

Europäisches Patentamt European Patent Office

Office européen des brevets



(11)

EUROPEAN PATENT APPLICATION

(12)

and the state of the state of

- (43) Date of publication: 01.10.1997 Bulletin 1997/40
- (51) Int CI 6 C12N 15/12. C12N 15/62, C12N 15/85, C07K 14/72, C12N 1/21, C12N 5/16, C12Q 1/00, C12Q 1/68
- (21) Application number: 97200903.9
- (22) Date of filing: 25.03.1997
- AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC (84) Designated Contracting States:
- NL PT SE (30) Priority 22.11.1996 EP 96203284 26.03.1996 EP 96200820
- (71) Applicant: Akzo Nobel N.V. 6824 BM Arnhem (NL)
- (72) Inventors: · Mosselman, Sietse 5346 VM OSS (NL)

- . Dijkema, Rein 5345 ML Oss (NL)
- Ogilvie-Emanuelson, Claudia Maria et al (74) Representative: Patent Department Pharma N.V. Organon P.O. Box 20 5340 BH Oss (NL)

- Estrogen receptor
- The present invention relates to isolated DNA encoding novel estrogen receptors, the proteins encod-(54)

ed by said DNA, chimeric receptors comprising parts of said novel receptors and uses thereof.

Printed by Jouve, 75001 PARIS (FR)

Description

30

40

100 The second

This invention relates to the field of receptors belonging to the superfamily of nuclear hormone receptors, in particular to steroid receptors. The invention relates to DNA encoding a novel steroid receptor, the preparation of said receptor, the recentor protein, and the uses thereof

epior. The receptor protein, and the uses thereof.

Steroid hormone receptors belong to a superfamily of nuclear hormone receptors involved in ligand-dependent steroid hormone receptors belong to a superfamily of nuclear hormone receptors for non-elegated hormone receptors belong to a superfamily consists of receivers for non-elegated hormone receptors and the cure demine the cure demand on the cure d Teliore normalite receptors belong to a supertamity of nuclear normalite receptors for non-steroid harmones transcriptional control of gene expression. In addition, this superfamily consists of receptors for non-steroid harmones and religions of all Nature 20, 624-620, 1087. Evans R.M. Science such as utamine D. thursid harmones and religions of the superfamily control of gene expression. such as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine U, thyroid hormonas and retinoids (Giguere et al. Nature 330, 624-629, 1987; Evans, H.M., Science socialized as vitamine up to the vitamine up receptor, the receptor protein, and the uses thereof. 240. 389-895, 1988). Moreover, a range of nuclear receptor-like sequences have been identified which encode socialed forth and therefore classified as nuclear receptors, although orphan' receptors: these receptors are structurally related to and therefore classified as nuclear receptors. I Manneterior orphan' receptors: these receptors are structurally related to and therefore classified as nuclear receptors. I Manneterior orphan' receptors: these receptors are structurally related to and therefore classified as nuclear receptors. In the control of the co orphan receptors: these receptors are structurally related to and ineretore classified as nuclear receptors, authority or putative ligands have been identified yet (B.W. O'Malley, Endocrinology 125, 1119-1170, 1989; D.J. Mangelsdorf and P.M. Euges Coll. 93, 241-250, 1005.

I.N.N. Evens, Oen, 93, 941-930, 1883.

The superfamily of nuclear hormone receptors share a modular structure in which six distinct structural and functions that the superfamily of nuclear hormone receptors share a modular structure in which six distinct structural and functions.

The superfamily of nuclear hormone receptors share a modular structure in which six distinct structural and functions.

The superfamily of nuclear hormone receptors share a modular structure in which six distinct structural and functions. The superfamily of nuclear hormone receptors share a modular structure in which six distinct structural and functional domains. A 10 F, are displayed (Evans, Science 240, 889-895, 1988). A nuclear hormone receptor is characterized tional domains. A 10 F, are displayed (Evans, Science 240, 889-895, 1988). In the superfamily of nuclear hormone receptors share a modular structure in which six distinct structure in tionaldomains, Ator, are displayed (Evans, Science 240, 889-895, 1988). A nuclear normone receptor is characterized by a variable N-terminal region (domain A/B), followed by a centrally located, highly conserved DNA-binding domain by a variable N-terminal region (domain A/B), a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a conserved linand-binding domain by a variable binder region (domain D) a variable binder r by a variable N-terminal region (domain A/b), followed by a centrally located, nightly conserved DINA-binding domain (thereinafter referred to as DBD; domain C), a variable hings region (domain D), a conserved ligand-binding domain (thereinafter referred to as DBD; domain E) and a variable C-terminal region (domain E) and R.M. Evans, Cell, <u>83</u>, 841-850, 1995).

telli aller referred to as LDU, domain E) and a variable U-terminal region (domain F).

The N-terminal region, which is highly variable in size and sequence, is poorly conserved among the different in the Modulation of transcription activation (Rocque) in the Modulation of transcription activation (Rocque) in the Modulation of transcription activation of the recent of the recent of the superfamily. This part of the recent of the Modulation of transcription activation (Rocque) (herein after referred to as LBD; domain E) and a variable C-terminal region (domain F). The IN-terminal region, which is highly variable in size and sequence, is poonly conserved among the different members of the superfamily. This pan of the receptor is involved in the modulation of transcription activation (Bocquel members of the superfamily. This pan of the receptor is involved in the modulation of transcription activation (Bocquel members of the superfamily. This pan of the receptor is involved in the modulation of transcription activation (Bocquel Machine Machine).

II, NUCL. ACID MES., 17, 2001-2000, 1909; lord et al., Cell 59, 477-487, 1989).

The DBD consists of approximately 66 to 70 amino acids and is responsible for DNA-binding activity: it targets the approximately 66 to 70 amino acids and is responsible for DNA-binding activity: it targets the Theology of the Acid Mes., 17, 2001-2000, 1909-1909, and the Acid Mes., 17, 2001-2000, as HARS) within the approximately 66 to 70 amino acids and is responsible for DNA-binding activity: it targets the acid Mes., 17, 2001-2000, as HARS) within the acid Mes., 17, 2001-2000, as HARS). The UBU consists of approximately be to 70 amino acids and is responsible for UNA-binding activity. It targets the receptor to specific DNA sequences called hormone responsive elements (hereinafter referred to as HRE) within the receptor to specific DNA sequences called hormone responsive elements (Martinez and Wahli In Nuclear Hormone Recentors). et al. Nucl. Acid Res., 17, 2581-2595, 1989; Tota et al., Cell 59, 477-487, 1989).

receptor to specific UNA sequences called hormone responsive elements (hereinalter reterred to as HHE) within the transcription control unit of specific target genes on the chromatin (Martinez and Wahli, In 'Nuclear Hormone Receptors', Acad. Press. 125-153. 1991).

10. Press, 123-133, 1331).

The LBD is located in the C-terminal part of the receptor and is primarily responsible for ligand binding activity. In the LBD is located in the C-terminal part of the hormone linand and in addition necessaries a transcring the LBD is occasional for recognition and hinding of the hormone linand and in addition necessaries. The LBU is located in the U-terminal part of the receptor and is primarily responsible for figand clinding activity. In this way, the LBD is essential for recognition and binding of the hormone figand and, in addition possesses a transcription are the specific transcription and the specific transcription activation function. Thereby determining the specificity and selectivity of the hormone response of the recents to activation function. this way, the LBD is essential for recognition and binding of the hormone ligand and, in addition possesses a transcription are supported by the hormone response of the receptor. It is essential for recognition and binding of the hormone response of the receptor. It is essential to recognition and binding of the hormone response of the receptor. It is essentially the specificity and selectivity of the hormone response of the receptor. It is essentially the specificity and selectivity of the hormone response of the receptor. It is essentially the selection of the hormone in addition possesses a transcription and binding of the hormone ligand and, in addition possesses a transcription and binding of the hormone response of the receptor. uon activation function, inereoy determining the specificity and selectivity of the normone response of the receptor.

Although moderately conserved in structure, the LBD's are known to vary considerably in homology between the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology between the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology between the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology between the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology between the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology between the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the LBD's are known to vary considerably in homology to the indiAlthough moderately conserved in structure, the latter than the indiAlthough moderately conserved in structure, the latter than the indiAlthough moderately conserved in structure, the latter than the indiAlthough moderately conserved in structure, the latter than the latter than the indiAlthough moderately conserved in structure, the latter than Although moderately conserved in structure, the LBU's are known to vary considerably in homology between the individual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB vidual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, F Acad. Press, 125-153, 1991). 2, JUNE JUNES, 1991, manyersouri et al. Cell, vol. 83, 833-839, 1995).

Functions present in the N-terminal region, LBD and DBD operate independently from each other and it has been already from the specific property of t 25

This results in chapter publisher recentors such as described for instance in MC-A-QORGES. Vidual Hieringers of the hucker Holling levels, oct. J. 5, 3092-3099, 1991; Mangelsdorf et al. Cell, Vol. 83, 835-839, 1995). when a hormone ligand for a nuclear receptor enters the cell by diffusion and is recognized by the LBD, it will bind when a hormone ligand for a nuclear receptor enters the cell by diffusion and is recognized by the LBD, it will bind to the specific recently protein the recent of t Shown that these domains can be exchanged between nuclear receptors (Green et al., Nature, This results in chimeric nuclear receptors, such as described for instance in WO-A-8905355.

when a normone liganulior a nuclear receptor enters the cell by diffusion and is recognized by the LDU, it will brid to the specific receptor protein, thereby initiating an allosteric alteration of the receptor protein, hereby initiating an allosteric alteration of the receptor protein, thereby initiating an allosteric alteration of the receptor protein, thereby initiating an allosteric alteration and as such is able to bind through the specific receptor protein, thereby initiating an allosteric alteration of the receptor protein. to the specific receptor protein, intereby initiating an allosteric alteration of the receptor protein. As a result of this able to bind through alteration the ligand/receptor complex switches to a transcriptionally active state and as such is able to bind through alteration the ligand/receptor complex switches to a transcriptionally active state and as such is able to bind through alteration the ligand/receptor complex switches to a transcriptionally active state and as such is able to bind through the presence of the DRD with high affinity to the corresponding HRF on the chromatic DNA (Martinez and Wahl). alleration the ligandreceptor complex switches to a transcriptionally active state and as such is able to bind through the presence of the DBD with high affinity to the corresponding HRE on the chromatin DNA (Martinez and Wahli, which is a presence of the DBD with high affinity to the corresponding the binandrecenter complex modulates as the presence of the DBD with high affinity to the corresponding the binandrecenter complex modulates as the presence of the DBD with high affinity to the corresponding the corr the presence of the UBU with high affinity to the corresponding that on the chromatin UNA (Matthez and virginity) to the corresponding that on the chromatin UNA (Matthez and virginity) to the chromatin UNA (Matthez and virginity) to the chromatin UNA (Matthez and virginity) to the chromatin under the chro Nuclear Hormone receptors, 123-133, Acad. Press, 1991). In this way the ligand/receptor complex modulates expression of the specific target genes. The diversity achieved by this family of receptors results from their ability to respond to different linands.

pond to different ligands.

The steroid hormone receptors are a distinct class of the nuclear receptor superfamily, characterized in that the interest of the nuclear receptor superfamily, characterized in that the interest ligands.

The steroid hormone is the recenters for altropoliticaids (ICR) mineral continuids (IMR) are needed to be a steroid hormone. The recenters for altropoliticaids (ICR) mineral continuids (IMR) are needed to be a steroid hormone. Ine steroid normone receptors are a distinct class of the nuclear receptor superiaritity, characterized in that the ligands are steroid hormones. The receptors for glucocorticoids (GR), mineralcorticoids (MR), progestins (PR), and or ligands are steroid hormones. The receptors for glucocorticoids (GR), mineralcorticoids (MR), progestins (PR), and or classical staroid recentors.

iliganus are sieroiu normones. The receptors ioi giucocorricolus (UR), mineralcorricolus (WR), progestins (PR), ariuogens (AR) and estrogens (ER) are classical steroid receptors. Furthermore, the steroid receptors The CR MR DR
aniity upon activation to hind to patindromic DNA sequences the socialized HRE's as homodimers. The CR MR DR gens (AH) and estrogens (EH) are classical steroid receptors. Furthermore, the steroid receptors have the unique ability upon activation to bind to palindromic DNA sequences, the so-called HRE's, as homodimers. The GR, MR, PR ability upon activation to bind to palindromic DNA sequence while the FR recognizes a different DNA sequence (Reado et al. Cell and AR recognize the same DNA sequence). ability upon activation to bind to patindromic DNA sequences, the so-called HHE's, as nomodimers. The GH, MH, PH and AR recognize the same DNA sequence. While the ER recognizes a different DNA sequence. (Beato et al., Cell, and AR recognize the same DNA sequence while the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the head of the started recentor is thought to interact with components of the started recentor is thought to interact with components of the started recentor is thought to interact with components of the started recentor is the started rece respond to different ligands. and AM recognize the same DNA sequence, while the EM recognizes a dillerent DNA sequence. (Deato et al., Cell, Vol. 83, 851-857, 1995). After binding to DNA, the steroid receptor is thought to interact with components of the basal vol. 83, 851-857, 1995). After binding to DNA, the steroid receptor is thought to interact with components of enacting transcription factors. Thus mortilation the expression of enacting transcription factors. VOI. 83, 851-857, 1995). After pinning to LINA, the steroid receptor is inought to interact with components of the basal transcriptional machinery and with sequence-specific transcription factors, thus modulating the expression of specific transcriptional machinery and with sequence-specific transcription.

get genes.

Several HRE's have been identified, which are responsive to the hormone/receptor complex. These HRE's are several HRE's have been identified, which are responsive to the hormone/receptor complex. These HRE's are several HRE's have been identified, which are responsive to the hormone/receptor complex. These HRE's are several HRE's have been identified, which are responsive to the hormone/receptor complex. These HRE's are several HRE's have been identified, which are responsive to the hormone/receptor complex.

Several MHE's nave been identified, which are responsive to the hormone/teceptor complex. These first started in the transcriptional control units of the various target genes such as mammalian growth hormone genes situated in the transcriptional control units of the various target genes and progestering received to observe the observed as the control of the various target genes and progestering received to observe the observed to observe the observe the observed to observe the observe the observed to obse situated in the transcriptional control units of the various target genes such as mammalian growth normone genes (responsive to glucocorticoid, estrogen, testosterone), mammalian prolactin genes and progesterone receptor genes (responsive to glucocorticoid, estrogen, testosterone), mammalian prolacting mammalian metallicitionein nene (re-(responsive to glucocorticold, estrogen, testosterone), mammalian protactin genes and progesterone receptor genes (responsive to progesterone), mammalian metallothionein gene (responsive to progesterone), mammalian metallothionein gene (responsive to progesterone), mammalian metallothionein gene (responsive to progesterone), avian ovalbumin genes (responsive to progesterone), avian ovalbumin genes (responsive to progesterone), avian ovalbumin genes (responsive to progesterone). (responsive to Estrogen), avian ovaibumin genes (responsive to progesterone), mammalian metallothionein gene (responsive to estrogen, tostosterone, glucosponsive to glucocorticoid) and mammalian hopatic $\omega_{2\mu}$ -globulin gene (responsive to estrogen, tostosterone, glucocorticoid). iculu).

The steroid hormone receptors have been known to be involved in embryonic development, adult homeostasis as

It as arread abusinlant. Marious dispasse and abnormalising house has associated to a disturbance in the staroid hor.

The steroid normone receptors have over known to be involved in embryonic development, adult normouslasts as well as organ physiology. Various diseases and abnormalities have been ascribed to a disturbance in the steroid moderate of their influence as hormone adiabated transmissional modulators. well as organ physiology. Various diseases and apnormalilies have been ascribed to a disturbance in the steroid northway. Since the steroid receptors exercise their influence as hormone-activated transcriptional modulators, and he anticinated that mutations and defects in these recentors as well as overetimitation or blocking of these if these recentors as well as overetimitation or blocking of these in these recentors. mone painway. Since the steroic receptors exercise their influence as normone-activated transcriptional modulators, it can be anticipated that mutations and defects in these receptors, as well as overstimulation or blocking of these it can be anticipated that mutations and defects in these receptors. receptors might be the underlying reason for the altered pattern. A better knowledge of these receptors, their mecha-nerm of action and of the linancis which hind to said recentor might help to create a hatter insight in the underlying receptors might be the underlying reason for the altered pattern. A better knowledge of these receptors, their mechanism of action and of the ligands which bind to said receptor might help to create a better insight in the underlying nism of action and of the ligands which bind to said receptor might help to create a better insight in the underlying nism of action and of the hormone signal transduction nathway which eventually will lead to hetter treatment of the disease. nism of action and of the ligands which bind to said receptor might help to create a better insight in the underlying mechanism of the hormone signal transduction pathway, which eventually will lead to better treatment of the diseases mechanism of the hormone signal transduction pathway.

abnormalities linked to altered normone/receptor functioning.

For this reason cDNA's of the steroid and several other nuclear receptors of several mammalians, including huse for this reason cDNA's of the steroid and several other nuclear receptors of several mammalians, including huse for this reason cDNA's of the steroid and several other nuclear receptors of several mammalians. mans, have been isolated and the corresponding amino acid sequences have been deduced, such as for example the human steroid receptors for vitamine D, thyroid hormones, human steroid receptors PR, ER, GR, MR, and AR, the human non-steroid receptors for vitamine D, thyroid hormones, human steroid receptors PR, ER, GR, MR, and AR, the human non-steroid receptors PR, ER, GR, MR, and relinning acid. In addition, child's anording well over 100 mammalian organization and relinning acid. In addition, child's anording well over 100 mammalian organization. and abnormalities linked to altered hormone/receptor functioning. human steroid receptors PH, EH, GH, MH, and AH, the human non-steroid receptors for vitamine D, thyroid hormones, and retinoids such as retinol A and retinoids acid. In addition, cDNA's encoding well over 100 mammalian orphan real and retinoids such as retinol A and retinoids acid. In addition, cDNA's encoding well over 100 mammalian orphan replacements as a language of the second of the sec and retinoids such as retinol A and retinoic acid. In addition, cDNA's encoding well over 100 mammalian orphan receptors have been isolated, for which no putative ligands are known yet (Mangelsdorf et al. Cell, Vol.83, 835-839,
ceptors have been isolated, for which no putative ligands are known yet of the elucidation of other nuclear recentors in order to unravet the various
the provided However there is still a great need for the elucidation of other nuclear recentors in order to unravet there is still a great need for the elucidation of other nuclear recentors. ceptors have been isolated, for which no putative liganus are known yet (wangetsoon et al. Cell, vol.33, 300-334, 1995). However, there is still a great need for the elucidation of other nuclear receptors in order to unravel the various roles those receptors play in normal objections and national collections.

1/2/ 1/2 / 1

15

20

30

45

The second

is these receptors play in normal physiology and pathology.

The present invention provides for such a novel nuclear receptor. More specific, the present invention provides for such a novel nuclear receptor. More specific, the present invention provides for such a novel nuclear receptor. More specific, the present invention provides for such a novel nuclear receptor. The present invention provides for such a novel nuclear receptor. More specific, the present invention provides for such a novel nuclear receptor. More specific, the present invention provides for such a novel steroid receptors are novel estrogen receptor novel steroid receptors. having estrogen mediated activity. Said novel steroid receptors are novel estrogen mediated activity. Said novel steroid receptors are novel estrogen mediated activity. Said novel steroid receptors are novel estrogen receptors, which are able to bind and be activated by for example, estradiol, estrone and estrictions are novel estrogen receptors. roles these receptors play in normal physiology and pathology. which are able to bind and be activated by, for example, estradiol, estrone and estriol.

According to the present invention it has been found that a novel estrogen receptor is expressed as an 8 kb.

According to the present invention it has been found that a novel estrogen receptor.

According to the present invented and parishers blood timeshouter (pare) over and tention. tor novel steroid receptors, naving estrogen mediated activity. Said novel steroid teceptors and estroid tors, which are able to bind and be activated by, for example, estradiol, estrone and estroid tors, which are able to bind and be activated if her hear found that a point estrone recent

According to the present invention it has been found that a novel estrogen receptor is expressed as an 8 kb transcript in human thymus, spleen, peripheral blood lymphocytes (PBLs), ovary and testis. Furthermore, and spleen transcript in human thymus, spleen, peripheral blood lymphocytes (PBLs), ovary and testis. Furthermore, and spleen transcript of anoroximately 10 kb was identified in ovary thymus and spleen transcript of anoroximately 10 kb was identified in ovary thymus and spleen transcript of anoroximately 10 kb was identified another transcript of anoroximately 10 kb was identified anoroximately 10 kb was identified anoroximately 10 kb was identified anoroximatel transcript in numan triymus, spieen, peripheral blood lymphocytes (PBLs), ovary and testis. Furthermore, additional transcript in numan triymus, spieen, peripheral blood lymphocytes (PBLs), ovary and testis. Furthermore, additional transcript in numan triymus, spieen, peripheral blood lymphocytes (PBLs), ovary and testis. Furthermore, additional transcript of approximately 10 kb was identified in ovary, thymus and spiecific transcripts are nonably denoted by alternative splicing. These transcripts are nonably denoted by alternative splicing in testis. an additional transcript of 1.3 kb was identified. transcripts have been identified. Another transcript of approximately 10 kb was identified in ovary, thymus and spiech.

Intestis, an additional transcript of 1.3 kb was detected. These transcripts are probably generated by alternative splicing of the gene ancording the power estimate according to the invention. ne gene encoding the novel estrogen receptor according to the invention, the invention revealed that several according to the invention revealed that several coloring of the cDNA's encoding the novel estrogen receptors according to the invention revealed that charminal coloring of the cDNA's encoding the novel estrogen receptors according to the order than the coloring that the color is the color of the cDNA's encoding the novel estrogen receptors according to the color of the cDNA's encoding the novel estrogen receptors according to the invention revealed that several colors according to the color of the cDNA's encoding the novel estrogen receptors according to the invention revealed that several colors according to the col of the gene encoding the novel estrogen receptor according to the invention.

Cloning of the cDNA's encoding the novel estrogen receptors according to the invention revealed that several splicing variants of said receptor can be distinguished. At the protein level, these variants differ only at the C-terminal splicing variants of said receptor can be distinguished. CDNA encoding an ER has been isolated (Green, et al, Nature 320, 134-139, 1986; Greene et al, Science 231, or contract the recent of the recen part.

cuna encoding an EH has been isolated (Green, et al., Nature 320, 134-139, 1986; Greene et al., Science 231, 1150-1154, 1986), and the corresponding amino acid sequence has been deduced. This receptor and the receptor acid sequence has been deduced. This receptor and the receptor acid sequence has been deduced. This receptor and the receptor acid sequence has been deduced. This receptor acid different sequences with different nucleic acid and encoded for his different sequences with different nucleic acid accounts to the present invention. 1150-1154, 1986), and the corresponding amino acid sequence has been deduced. This receptor and the receptor according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid according to the present invention. according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid to the ER according to the ER according to the sequences. Not only do the ER of the prior art (hereinafter referred to as classical ER) and the ER according to the sequences. Not only do the ER of the prior art (hereinafter referred to as classical ER) and the ER according to the sequences. Not only do the ER of the prior art (hereinafter referred to as classical ER) and the ER according to the sequences. sequences. Not only do the EH of the prior art [hereinatter referred to as classical EH] and the EH according to the prior art [hereinatter referred to as classical EH] and the EH according to the encoding the EH of the prior art [hereinatter referred to as classical EH] and the EH according to the invention was found the classical EH is located on chromosome 6 whereas the gene encoding the ER according to the invention was found to the classical EH is located on chromosome 6 whereas the gene encoding the ER according to the invention was found to the invention was found to the classical EH is located on chromosome 6. present invention differ in amino acid sequence, they also are located on different chromosomes. The gene encounty the classical ER is located on chromosome 6, whereas the gene encoding the ER according to the invention the classical transfer distinguishes itself from the classical to be located on chromosome 14. The ER according to the invention furthermore distinguishes itself from the classical transfer distinguishes and the classical transfer distinguishes distinguished transfer distinguishes distinguished transfer distinguishes distinguished transfer distinguished transfer distinguishes distinguished transfer distinguishe the classical ER is located on chromosome 6, whereas the gene encoding the ER according to the invention was found to be located on chromosome 14. The ER according to the invention furthermore distinguishes itself from the classical to be located on chromosome 14. The ER according to the invention furthermore hatween these recentors in differences in tissue distribution. to be located on chromosome 14. The EH according to the invention turnermore distinguishes lisely from the classical receptor in differences in tissue distribution, indicating that there may be important differences between these receptors at the level of detropodic cinnalling.

he level of estrogenic signalling.

In addition, two orphan receptors, ERRa and ERRB, having an estrogen receptor related structure have been hardling, two orphan receptors, ERRa and ERRB, having an estrogen receptor have not been renorted to be actived forward at al. Nature and an analyst and these orphan recentors. In addition, two orphan receptors, EHHα and EHHβ, having an estrogen receptor related structure have been described (Giguere et al., Nature 331, 91-94, 1988). These orphan receptors, however, have not been reported to be described (Giguere et al., Nature 331, 91-94, 1988). These orphan receptors, however, have not been reported to the classical ER and other linands which hind to the classical ER and other linands which hind to the classical ER. described (Giguere et al., Nature 331, 91-94, 1988). These orphan receptors, however, have not been reported to be able to bind estrodial or any other hormone that binds to the classical ER, and other ligands which bind to these able to bind estrodial or any other hormone that binds to the classical in the invention distinguishes itself classical to the classical estrodiant to the invention distinguishes itself classical estrodiants. able to bind estrodial or any other hormone that binds to the classical EH, and other ligands which bind to these receptors have not been found yet. The novel estrogen receptor according to the invention distinguishes itself clearly from these recentors since it was found to hind estronens. at the level of estrogenic signalling.

If these receptors since it was round to bind estrogens.

The fact that a novel ER according to the invention has been found is all the more surprising, since any suggestion of large that a novel end according to the invention was absent in the scientific literature, neither the isolation of large that a novel end according to the invention was absent in the scientific literature, neither the isolation of large that a novel end according to the invention was absent in the scientific literature. The fact that a novel EH according to the invention has been found is all the more surprising, since any suggestion towards the existence of additional estrogen receptors was absent in the scientific literature: neither the isolation of towards the existence of additional estrogen receptors was absent in the scientific literature: neither the oresence of additional estrogen receptors was absent in the scientific literature: neither the oresence of additional estrogen receptors. towards the existence of additional estrogen receptors was absent in the scientific interactive: retirner the isolation of the classical ER nor the orphan receptors ERRic and ERRig suggested or hinted towards the presence of additional the classical ER nor the orphan receptors ERRic and ERRig suggested or hinted towards the presence of additional ER's could be a classical ER nor the orphan receptors according to the invention. The identification of additional ER's could be according to the invention. from these receptors since it was found to bind estrogens. the classical EH nor the orphan receptors EHHa and EHHb suggested or hinted towards the presence of additional EH's could be a estrogen receptors such as the receptors according to the invention. The identification of additional EH's could be a estrogen receptors such as the receptors according to the invention. The identification of one FR and as such ascribe estrogen receptors such as the receptors according to the invention. estrogen receptors such as the receptors according to the invertion. The identification of additional Ensignment as such ascribe major step forward for the existing clinical therapies, which are based on the existence of one ER and as such ascribe major step forward for the existing clinical therapies, which are based on the existence according to the invention will all setrogen mediated abnormalities and/or dispasse to this one recentor. major step forward for the existing clinical therapies, which are based on the existence of one ER and as such ascribe all estrogen mediated abnormalities and/or diseases to this one receptor. The receptors according to the invention will estrogen mediated abnormalities and/or diseases to this one receptor. The receptors according to the invention will estrogen mediated abnormalities and/or diseases to this one receptor. The receptors according to the invention will be development of hormone analone that selectively activate either the classical ER or the novel estrogen that selectively activate either the classical ER or the novel estrogen. all estrogen mediated abnormalities and/or diseases to this one receptor. The receptors according to the invention will be useful in the development of hormone analogs that selectively activate either the classical ER or the novel estrogen be useful in the development of hormone analogs that selectively activate either the classical ER or the novel estrogen be useful in the development of hormone analogs that selectively activate either the classical ER or the novel estrogen because of the national development of hormone analogs that selectively activate either the classical ER or the novel estrogen because of the national development of hormone analogs that selectively activate either the classical ER or the novel estrogen because of the national development of hormone analogs that selectively activate either the classical ER or the novel estrogen because of the national development of hormone analogs that selectively activate either the classical ER or the novel estrogen because of the national development of hormone analogs that selectively activate either the classical ER or the national development of hormone analogs that selectively activate either the classical end of the national development of hormone analogs that selectively activate either the classical end of the national development o be useful in the development of hormone analogs that selectively activate either the classical ER or the novel estrogen receptor according to the invention. This should be considered as one of the major advantages of the present invention. This is one aspect, the present invention provides for replaced above a provide entering a novel eternic repeator. spior according to the invention. This should be considered as one of the major advantages of the present invention.

Thus, in one aspect, the present invention provides for isolated cDNA encoding a novel steroid receptor. In particular, the present invention provides do not provide a possible forms.

lar, the present invention provides for isolated cunk encoding a novel estrogen receptor.

According to this aspect of the present invention, there is provided an isolated DNA encoding a steroid receptor and a translation domain, wherein the amino acid according to this aspect of the present invention, there is provided an isolated DNA encoding the amino acid according to this aspect of the present invention, there is provided an isolated DNA encoding the amino acid. ticular, the present invention provides for isolated cUNA encoding a novel estrogen receptor, ticular, the present invention provides for isolated cDNA encoding a novel estrogen receptor. According to this aspect of the present invention, there is provided an isolated DNA encoding a steroid receptor protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid protein having an N-terminal domain, a DNA-binding domain and a ligand-binding homology with the amino acid sequence of said DNA-binding domain of said recently orders exhibits at least 20% homology with the amino acid

protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said receptor protein exhibits at least 80% homology with the amino acid sequence of said linand-hinding domain of said recentor protein exhibits at least 80% homology with the amino acid sequence of said linand-hinding domain of said recentor protein exhibits at least 80% homology with the amino acid sequence of said linand-hinding domain of said recentor protein exhibits at least 80% homology with the amino acid sequence of said linand-hinding domain, wherein the amino acid

sequence of said DNA-binding domain of said receptor protein exhibits at least 80% homology with the amino acid sequence of said ligand-binding domain of said receptor protein exhibits at least 70%, homology with the amino acid sequence of said ligand-binding domain of said receptor protein sequence shown in SEO ID NO.4 nions at least 70% nomology with the amino acid sequence shown in SEQ 10 NO.4.

In particular, the isolated DNA encodes a steroid receptor protein having an N-terminal domain, a DNA-binding In particular, the isolated DNA encodes a steroid receptor protein having an NA-binding domain whorein the amino acid sequence of said DNA-binding domain whorein the amino acid sequence of said DNA-binding domain. sequence shown in SEQ ID NO.4. and the amino acid sequence of salo ligand-ornaring or exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO.4. In particular, the isolated UNA encodes a steroid receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain of said receptor protein h domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said receptor protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence exhibits at least 90%, preferably 95%, more preferably 98%, m

d sequence snown in SEC ID NO.3.

More particularly, the isolated DNA encodes a steroid receptor protein having an N-terminal domain of said recentor.

More particularly, the isolated DNA encodes a steroid receptor protein having an N-terminal domain of said recentor. More particularly, the isolated DNA encodes a steroid receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain, a DNA-binding domain of said receptor protein having an N-terminal domain of said receptor prote domain and a ligand-binding domain, wherein the amino acid sequence of said ligand-binding domain of said receptor of said ligand-binding domain, wherein the amino acid sequence of said ligand-binding domain, wherein the amino acid sequence of said ligand-binding domain of said receptor of said ligand-binding domain of acid sequence shown in SEQ ID NO.3.

acid sequence shown in SEQ ID NO:4.

35

45

 $(e_1,e_2)_{1\leq i\leq k} \in \mathcal{C}^{n_i}$

William Company

Berling to the training of the said

A preferred isolated DNA according to the invention encodes a steroid receptor protein having the amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.

A more preferred isolated DNA according to the invention is an isolated DNA comprising a nucleotide sequence stown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20 or SEQ ID NO:24.

The DNA according to the invention may be obtained from cDNA. Alternatively, the coding sequence might be genomic DNA or prepared using DNA synthesis techniques.

The DNA according to the invention will be very useful for in vivo expression of the novel receptor proteins according to the invention in sufficient quantities and in substantially pure form.

In another aspect of the invention, there is provided for a steroid receptor comprising the amino acid sequence encoded by the above described DNA molecules.

The steroid receptor according to the invention has an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said receptor exhibits at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and the amino acid sequence of said ligand-binding domain of said receptor exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4.

In particular, the steroid receptor according to the invention has an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said receptor exhibits at least 90% preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence shown in SEC ID NO 3.

More particular, the steroid receptor according to the invention has an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said ligand-binding domain of said receptor exhibits at least 75% prefearbly 80%, more preferably 90%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO 4

It will be clear for those skilled in the art that also steroid receptor proteins comprising combined DBD and LBD preferences and DNA encoding such receptors are subject of the invention.

Preferably, the steroid receptor according to the invention comprises an amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.

Also within the scope of the present invention are steroid receptor proteins which comprise variations in the amino acid sequence of the DBD and LBD without loosing their respective DNA-binding or ligand-binding activities. The variations that can occur in those amino acid sequences comprise deletions, substitutions, insertions, inversions or additions of (an) amino acid(s) in said sequence, said variations resulting in amino acid difference(s) in the overall sequence. It is well known in the art of proteins and peptides that these amino acid differences lead to amino acid sequences that are different from, but still homologous with the native amino acid sequence they have been derived from.

Amino acid substitutions that are expected not to essentially alter biological and immunological activities, have been described in for example Dayhof, M.D., Atlas of protein sequence and structure, Nat. Biomed. Res. Found., Washington D.C., 1978, vol. 5, suppl. 3. Amino acid replacements between related amino acids or replacements which have occurred frequently in evolution are, inter alia Ser/Ala, Ser/Gly, Asp/Gly, Arg/Lys, Asp/Asn, Ile/Val. Based on this information Lipman and Pearson developed a method for rapid and sensitive protein comparison (Science 227, 1435-1441, 1985) and determining the functional similarity between homologous polypeptides.

Variations in amino acid sequence of the DBD according to the invention resulting in an amino acid sequence that has at least 80% homology with the sequence of SEQ ID NO:3 will lead to receptors still having sufficient DNA binding activity. Variations in amino acid sequence of the LBD according to the invention resulting in an amino acid sequence that has at least 70% homology with the sequence of SEQ ID NO:4 will lead to receptors still having sufficient ligand highlight activity.

Homology as defined herein is expressed in percentages, determined via PCGENE. Homology is calculated as the percentage of identical residues in an alignment with the sequence according to the invention. Gaps are allowed to obtain maximum alignment.

Comparing the amino acid sequences of the classical ER and the ER's according to the invention revealed a high degree of similarity within their respective DBD's. The conservation of the P-box (amino acids E-G-X-X-A) which is responsible for the actual interactions of the classical ER with the target DNA element (Zilliacus et al., Mol.Endo. 9, 389, 1995; Glass. End.Rev. 15, 391, 1994), is indicative for a recognition of estrogen responsive elements (ERE's) by the ER's according to the invention. The receptors according to the invention indeed showed ligand-dependent transactivation on ERE-containing reporter constructs. Therefore, the classical ER and the novel ER's according to the invention may have overlapping target gene specificities. This could indicate that in tissues which co-express both respective ER's, these receptors compete for ERE's. The ER's according to the invention may regulate transcription of target genes differently from classical ER regulation or could simply block classical ER functioning by occupying estrogen responsive elements. Alternatively, transcription might be influenced by heterodimerization of the different

receptors.

يجيئ فالخرأة وتحمياهي الأخماء والوجيج

Thus, a preferred steroid receptor according to the invention comprises the amino acid sequence E-G-X-X-A within the P box of the DNA binding domain, wherein X stands for any amino acid. Also within the scope of the invention is isolated DNA encoding such a receptor.

Methods to prepare the receptors according to the invention are well known in the art (Sambrook et al., Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press. Cold Spring Harbor, 1989). The most practical approach is to produce these receptors by expression of the DNA encoding the desired protein.

A wide variety of host cell and cloning vehicle combinations may be usefully employed in cloning the nucleic acid sequence coding for the receptor of the invention. For example, useful cloning vehicles may include chromosomal non-chromosomal and synthetic DNA sequences such as various known bacterial plasmids and wider host range plasmids and vectors derived from combinations of plasmids and phage or virus DNA. Useful hosts may include bacterial hosts, yeasts and other fungi, plant or animal hosts, such as Chinese Hamster Ovary (CHO) cells or monkey cells and other hosts.

Vehicles for use in expression of the ligand-binding domain of the present invention will further comprise control sequences operably linked to the nucleic acid sequence coding for the ligand-binding domain. Such control sequences generally comprise a promoter sequence and sequences which regulate and/or enhance expression levels. Furthermore an origin of replication and/or a dominant selection marker are often present in such vehicles. Of course control and other sequences can vary depending on the host cell selected.

Techniques for transforming or transfecting host cells are quite known in the art (see, for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, 1989).

Recombinant expression vectors comprising the DNA of the invention as well as cells transformed with said DNA or said expression vector also form part of the present invention.

In a further aspect of the invention, there is provided for a chimeric receptor protein having an N-terminal domain, a DNA-binding domain, and a ligand-binding domain, characterized in that at least one of the domains originates from a receptor protein according to the invention, and at least one of the other domains of said chimeric protein originates from another receptor protein from the nuclear receptor superfamily, provided that the DNA-binding domain and the ligand-binding domain of said chimeric receptor protein originate from different proteins.

In particular, the chimeric receptor according to the invention comprises the LBD according to the invention, said LBD having an amino acid sequence which exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4. In that case the N-terminal domain and DBD should be derived from another nuclear receptor, such as for example PR. In this way a chimeric receptor is constructed which is activated by a ligand of the ER according to the invention and which targets a gene under control of a progesterone responsive element. The chimeric receptors having a LBD according to the invention are useful for the screening of compounds to identify novel ligands or hormone analogs which are able to activate an ER according to the invention.

In addition, chimeric receptors comprising a DBD according to the invention, said DBD having an amino acid sequence exhibiting at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and a LBD and, optionally, an N-terminal domain derived from another nuclear receptor, can be succesfully used to identify novel ligands or hormone analogs for said nuclear receptors. Such chimeric receptors are especially useful for the identification of the respective ligands of orphan receptors.

Since steroid receptors have three domains with different functions, which are more or less independent, it is possible that all three functional domains have been derived from different members of the steroid receptor superfamily.

Molecules which contain parts having a different origin are called chimeric. Such a chimeric receptor comprising the ligand-binding domain and/or the DNA-binding domain of the invention may be produced by chemical linkage, but most preferably the coupling is accomplished at the DNA level with standard molecular biological methods by fusing the nucleic acid sequences encoding the necessary steroid receptor domains. Hence, DNA encoding the chimeric receptor proteins according to the invention are also subject of the present invention.

Such chimeric proteins can be prepared by transfecting DHA encoding these chimeric receptor proteins to suitable host cells and culturing these cells under suitable conditions.

It is extremely practical if, next to the information for the expression of the steroid receptor, also the host cell is transformed or transfected with a vector which carries the information for a reporter molecule. Such a vector coding for a reporter molecule is characterized by having a promoter sequence containing one or more hormone responsive elements (HRE) functionally linked to an operative reporter gene. Such a HRE is the DNA target of the activated steroid receptor and, as a consequence, it enhances the transcription of the DNA, coding for the reporter molecule. In *in vivo* settings of steroid receptors the reporter molecule comprises the cellular response to the stimulation of the ligand. However, it is possible *in vitro* to combine the ligand-binding domain of a receptor to the DNA binding domain and transcription activating domain of other steroid receptors, thereby enabling the use of other HRE and reporter molecule systems. One such a system is established by a HRE presented in the MMTV-LTR (mouse mammary tumor virus long terminal repeat sequence in connection with a reporter molecule like the firefly luciferase gene or the bacterial gene

for CAT (chloramphenicol transferase). Other HRE's which can be used are the rat oxytocin promotor, the retinoic acid responsive element, the thyroid hormone responsive element, the estrogen responsive element and also synthetic responsive elements have been described (for instance in Fuller, ibid. page 3096). As reporter molecules next to CAT and luciferase β -galactosidase can be used.

Steroid hormone receptors and chimeric receptors according to the present invention can be used for the *in vitro* identification of novel ligands or hormonal analogs. For this purpose binding studies can be performed with cells transformed with DNA according to the invention or an expression vector comprising DNA according to the invention, said cells expressing the steroid receptors or chimeric receptors according to the invention.

The novel steroid hormone receptor and chimeric receptors according to the invention as well as the ligand-binding domain of the invention, can be used in an assay for the identification of functional ligands or hormone analogs for the nuclear receptors.

Thus, the present invention provides for a method for identifying functional ligands for the steroid receptors and chimeric receptors according to the invention, said method comprising the steps of

- a) introducing into a suitable host cell 1) DNA or an expression vector according to the invention, and 2) a suitable reporter gene functionally linked to an operative hormone response element, said HRE being able to be activated by the DNA-binding domain of the receptor protein encoded by said DNA;
 - b) bringing the host cell from step a) into contact with potential ligands which will possibly bind to the ligand-binding domain of the receptor protein encoded by said DNA from step a);
 - c) monitoring the expression of the receptor protein encoded by said reporter gene of step a).

If expression of the reporter gene is induced with respect to basic expression (without ligand), the functional ligand can be considered as an agonist; if expression of the reporter gene remains unchanged or is reduced with respect to basic expression, the functional ligand can be a suitable (partial) antagonist.

For performing such kind of investigations host cells which have been transformed or transfected with both a vector encoding a functional steroid receptor and a vector having the information for a hormone responsive element and a connected reporter molecule are cultured in a suitable medium. After addition of a suitable ligand, which will activate the receptor the production of the reporter molecule will be enhanced, which production simply can be determined by assays having a sensitivity for the reporter molecule. See for instance WO-A-8803168. Assays with known steroid receptors have been described (for instance S. Tsai et al., Cell 57, 443, 1989; M. Meyer et al., Cell 57, 433, 1989).

Legends to the figures

Figure 1.

5

20

25

35

45

55

Northern analysis of the novel estrogen receptor (ER β). Two different multiple tissue Northern blots (Clontech) were hybridised with a specific probe for ER β (see examples). Indicated are the human tissues the RNA originated from and the position of the size markers in kilobases (kb).

Figure 2.

Histogram showing the 3- to 4-fold stimulatory effect of 17β -estradiol, estriol and estrone on the luciferase activity mediated by En β . An expression vector encoding ER β was transiently transfected into CHO cells together with a reporter construct containing the rat oxytocin promoter in front of the firefly luciferase encoding sequence (see examples).

Figure 3.

Effect of 17 β -estradiol (E2) alone or in combination with the anti-estrogen ICI-164384 (ICI) on ER α and ER β . Expression constructs for ER α (the classical ER) and ER β were transiently transfected into CHO cells together with the rat oxytocin promoter-luciferase reporter construct described in the examples. Luciferase activities were determined in triplicate and normalised for transfection efficiency by measuring β -galactosidase in the same lysate.

Figure 4.

Expression of ERα and ERβ in a number of cell lines determined by RT-PCR analysis (see examples). The cell lines used were derived from different tissues/cell types: endometrium (ECC1, Ishikawa, HEC-1A, RL95-2); osteosarcoma (SAOS-2, U2-OS, HOS, MG63); breast tumours (MCF-7, T47D), endothelium (HUV-EC-C, BAEC-1); smooth muscle (HISM, PAC-1, A7R5, A10, RASMC, CavaSMC); liver (HepG2); colon (CaCo2); and vagina (Hs-760T, SW-954).

All cell lines were human except for PAC-1, A7R5, A10 and RASMC which are of rat origin, BAEC-1 which is of bovine origin and CavaSMC which is of guinea pig origin

Figure 5.

Sand Straight .

Transactivation assay using stably transfected CHO cell lines expressing ER α or ER β together with the rat oxytocin-luciferase estrogen-responsive reporter (see examples for details). Hormone-dependent transactivation curves were determined for 17 β -estradiol and for Org4094. For the ER antagonist raloxifen, cells were treated with 2 x 10⁻¹⁰ mol/L 17 β -estradiol together with increasing concentrations of raloxifen. Maximal values of the responses were arbitrarily set at 100%.

Examples

15

20

25

30

A. Molecular cloning of the novel estrogen receptor.

Two degenerate oligonucleotides containing inosines (I) were based on conserved regions of the DNA-binding domains and the ligand-binding domains of the human steroid hormone receptors.

Primer #1:

5'-GGIGA(C/T)GA(A/G)GC(A/T)TCIGGITG(C/T)CA(C/T)TA(C/T)GG-3'
(SEQ ID NO:7).

Primer #2:

5'-AAGCCTGG (C/G) A (C/T) IC (G/T) (C/T) TTIGCCCAI (C/T) TIAT-3' SEQ ID NO:8).

As template, cDNA from human EBV-stimulated PBLs (peripheral blood leukocytes) was used. One microgram of total RNA was reverse transcribed in a 20 µl reaction containing 50 mM KCl, 10 mM Tris-HCl pH 8.3, 4 mM MgCl2, 1 mM dNTPs (Pharmacia), 100 pmol random hexanucleotides (Pharmacia), 30 Units RNAse inhibitor (Pharmacia) and 200 Units M-MLV Reverse transcriptase (Gibco BRL). Reaction mixtures were incubated at 37°C for 30 minutes and heat-inactivated at 100°C for 5 minutes. The cDNA obtained was used in a 100 µl PCR reaction containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl2, 0.001% gelatin (w/v), 3% DMSO, 1 microgram of primer #1 and primer #2 and 2.5 Units of Amplitaq DNA polymerase (Perkin Elmer). PCR reactions were performed in the Perkin Elmer 9600 thermal cycler. The initial denaturation (4 minutes at 94°C) was followed by 35 cycles with the following conditions: 30 sec. 94°C, 30 sec. 45°C, 1 minute 72°C and after 7 minutes at 72°C the reactions were stored at 4°C. Aliquots of these reactions were analysed on a 1.5% agarose gel. Fragments of interest were cut out of the gel, reamplified using identical PCR-conditions and purified using Qiaex II (Qiagen). Fragments were cloned in the pCRII vector and transformed into bacteria using the TA-cloning kit (Invitrogen). Plasmid DNA was isolated for nucleotide sequence analysis using the Qiagen, plasmid midi protocol (Qiagen). Nucleotide sequence analysis was performed with the ALF automatic sequencer (Pharmacia) using a T7 DNA sequencing kit (Pharmacia) with vector-specific or fragment-specific primers.

One cloned fragment corresponded to a novel estrogen receptor (ER) which is closely related to the classical estrogen receptor. Part of the cloned novel estrogen receptor fragment (nucleotides 466 to 797 in SEQ ID 1) was amplified by PCR using oligonucleotide #3 TGTTACGAAGTGGGAATGGTGA (SEQ ID NO:9) and oligonucleotide #2 and used as a probe to screen a human testis cDNA library in \(\lambda\)gt11 (Clontech #HL1010b). Recombinant phages were plated (using Y1090 bacteria grown in LB medium supplemented with 0.2% maltose) at a density of 40.000 pfu (plaqueforming units) per 135 mm dish and replica filters (Hybond-N, Amersham) were made as described by the supplier. Filters were prehybridised in a solution containing 0.5 M phosphate buffer (pH 7.5) and 7% SDS at 65°C for at least 30 minutes. DNA probes were purified with Qiaex II (Qiagen), \(^32P\)-labeled with a Decaprime kit (Ambion) and added to the prehybridisation solution. Filters were hybridised at 65°C overnight and then washed in 0.5 X SSC/0,1% SDS at 65°C. Two positive plaques were identified and could be shown to be identical. These clones were purified by rescreening one more time. A PCR reaction on the phage eluates with the \(\lambda\)gt11-specific primers #4: 5'-TTGACAC-CAGACCAACTGGTAATG-3' (SEQ ID NO:10) and #5: 5'-GGTGGCGACGACTCCTGGAGCCCG-3' (SEQ ID NO:11)

yielded a fragment of 1700 basepairs on both clones.

Subsequent PCR reactions using combinations of a gene-specific primer #6: 5'-GTACACTGATTTGTAGCTGAC-3' (SEQ ID NO:12) with the \(\lambda \)glll primer #4 and gene-specific primer #7: 5'-CCATGATGATGTCCCTGACC-3' (SEQ ID NO:13) with \(\lambda \)gll1 primer primer #5 yielded fragments of approximately 450 bp and 1000 bp. respectively, which were cloned in the pCRII vector and used for nucleotide sequence analysis. The conditions for these PCR reactions were as described above except for the primer concentrations (200 ng of each primer) and the annealing temperature (60° C). Since in the cDNA clone the homology with the ER is lost abruptly at a site which corresponds to the exon 7/exon 8 boundary in the ER (between nucleotides 1247 and 1248 in SEQ ID NO:1), it was suggested that this sequence corresponds to intron 7 of the novel ER gene. For verification of the nucleotide sequences of this cDNA clone, a 1200 bp fragment was generated on the cDNA clone with \(\lambda \)gt11 primer #4 with a gene-specific primer #8 corresponding to the 3' end of exon 7: 5'-TCGCATGCCTGACGTGGGAC-3' (SEQ ID NO:14) using the proofreading \(Pfu \) polymerase (Stratagene). This fragment was also cloned in the pCRII vector and completely sequenced and was shown to be identical to the sequences obtained earlier.

To obtain nucleotide sequences of the novel ER downstream of exon 7, a degenerate oligonucleotide based on the AF-2 region of the classical ER (#9: 5'-GGC(C/G)TCCAGCATCTCCAG(C/G)A(A/G)GAG-3'; SEQ ID NO:15) was used together with the gene-specific oligonucleotide #10: 5'-GGAAGCTGGCTCACTTGCTG-3' (SEQ ID NO:16) using testis cDNA as template (Marathon ready testis cDNA, Clontech Cat #7414-1). A specific 220 bp fragment corresponding to nucleotides 1112 to 1332 in SEQ ID No. 1 was cloned and sequenced. Nucleotides 1112 to 1247 were identical to the corresponding sequence of the cDNA clone. The sequence downstream thereof is highly homologous with the corresponding region in the classical ER. In order to obtain sequences of the novel ER downstream of the AF-2 region, RACE (rapid amplification of cDNA ends) PCR reactions were performed using the Marathon-ready testis cDNA (Clontech) as template. The initial PCR was performed using oligonucleotide #11: 5'-TCTTGTTCTGGACAGGGATG-3' (SEQ ID NO:17) in combination with the AP1 primer provided in the kit. A nested PCR was performed on an aliquot of this reaction using oligonucleotide #10 (SEQ ID NO:16) in combination with the oligo dT primer provided in the kit. Subsequently, an aliquot of this reaction was used in a nested PCR using oligonucleotide #12: 5'-GCATGGAACATCTGCT-CAAC-3' (SEQ ID NO:18) in combination with the oligo dT primer. Nucleotide sequence analysis of a specific fragment that was obtained (corresponding to nucleotides 1256 to 1431 in SEQ ID NO 1) revealed a sequence encoding the carboxyterminus of the novel ER ligand-binding domain, including an F-domain and a translational stop codon and part of the 3' untranslated sequence which is not included in SEQ ID NO:1. The deduced amino acid sequence is shown in SEQ ID NO:5.

In order to investigate the possibility that the novel estrogen receptor had additional, upstream translation-initiation codons, RACE-PCR experiments were performed using Marathon-ready testis cDNA (Clontech Cat. # 7414-1). First a PCR was performed using oligonucleotide SEQ ID NO:12 (antisense corresponding to nucleotides 416-395 in SEQ ID NO:1) and AP-1 (provided in the kit). A nested PCR was then performed using oligonucleotide having SEQ ID NO:27 (antisense corresponding to nucleotides 254-231 in SEQ ID NO:1) with AP-2 (provided in the kit). From the smear that was obtained, the region corresponding to fragments larger than 300 basepairs was cut out, purified using the GenecleanlI kit (Bio101) and cloned using the TA-cloning kit (Clontech). Colonies were screened by PCR using genespecific primers: SEQ ID NO:22 and SEQ ID NO:28. The clone containing the largest insert was sequenced. The nucleotide sequence corresponds to nucleotides 1 to 490 in SEQ ID NO:24. It is clear from this sequence that the first in-frame upstream translation initiation codon is present at position 77-79 in SEQ ID NO:24. Upstream of this translational startcodon an in-frame stop-codon is present (11-13 in SEQ ID NO:24). Consequently, the reading frame of the novel estrogen receptor is 530 amino acids (shown in SEQ ID NO:25) and has a calculated molecular mass of 59.234

To contirm the nucleotide sequences obtained by 5' RACE, human genomic clones were obtained and analysed. A human genomic library in λEMBL3 (Clontech HL1067J) was screened with a probe corresponding to nucleotides 1 to 416 in SEQ ID NO:1. A strongly hybridizing clone was plaque-purified and DNA was isolated using standard protocols (Sambrook et al. 1989). The DNA was digested with several restriction enzymes, electrophoresed on agarose gel and blotted onto Nylon filters. Hybridisation of the blot with a probe corresponding to the above-mentioned RACE fragment (nucleotides 1-490 in SEQ ID NO:24) revealed a hybridizing Sau3A fragment of approximately 800 basepairs. This fragment was cloned into the BamH1 site of pGEM3Z and sequenced. The nucleotide sequence contained one base difference which is probably a PCR-induced point mutation in the RACE fragment. Nucleotide 172 was a G residue in the 5'RACE fragment, but an A residue in several independent genomic subclones.

B. Identification of two splice variants of the novel estrogen receptor.

55

Rescreening of the testis cDNA library with a probe corresponding to nucleotides 918 to 1246 in SEQ ID No. 1 yielded two hybridizing clones, the 3' end of which were amplified by PCR (gene-specific primer #10: 5'-GGAAGCT-GGCTCACTTGCTG-3' (SEQ ID NO:16) together with primer #4, SEQ ID NO:10), cloned and sequenced. One clone

was shown to contain an alternative exon 8 (exon 8B) of the novel ER. In SEQ ID No. 2 the protein encoding part and the stopcodon of this splice variant are presented. As a consequence of the introduction of this exon through an alternative splicing reaction, the reading frame encoding the novel ER is immediately terminated, thereby creating a trun-

Screening of a human thymus cDNA library (Clontech HL1074a) with the probe corresponding to nucleotides 918 cation of the carboxyterminus of the novel ER (SEQ ID NO:6). to 1246 in SEQ ID No. 1, revealed another splice variant. The 3' end of one hybridizing clone was amplified using primer #10 (SEQ ID NO:16) with the λgt10-specific primer #13.5'-AGCAAGTTCAGCCTGTTAAGT-3' (SEQ ID NO:19). cloned and sequenced. The obtained nucleotide sequence upstream of the exon 7/exon 8 boundary was identical to the clones identified earlier. However, an alternative exon 8 (exon 8C) was present at the 3' end encoding two Cterminal amino acids followed by a stop-codon. The nucleotide sequence of the protein-encoding part of this splice variant is shown in SEQ ID NO:20, the corresponding protein sequence is SEQ ID NO:21.

These two variants of the novel estrogen receptor do not contain the AF-2 region and therefore probably lack the ability to modulate transcription of target genes in a ligand-dependent fashion. However, the variants potentially could interfere with the functioning of the wild-type classical ER and/or the wild-type novel ER, either by heterodimerization or by occupying estrogen response elements or by interactions with other transcription factors. A mutant of the classical ER (ER1-530) has been described which closely resembles the two variants of the novel estrogen receptor described above. ER1-530 has been shown to behave as a dominant-negative receptor i.e. it can modulate the intracellular activity of the wild type ER (Ince et al, J. Biol. Chem. 268, 14026-14032, 1993).

and the second

المناعدية مهار

Human multiple tissue Northern blots (MTN-blots) were purchased from Clontech and prehybridized for at least 1 C. Northern blot analysis. hour at 65°C in 0.5 M phosphate buffer pH 7.5 with 7% SDS. The DNA fragment that was used as a probe (corresponding to nucleotides 466 to 797 in SEQ ID No. 1) was ³²P-labeled using a labelling kit (Ambion), denatured by boiling and added to the prehybridisation solution. Washing conditions were: 3X SSC at room temperature, followed by 3 X SSC at 65°C, and finally 1 X SSC at 65°C. The filters were than exposed to X-ray films for one week. Two transcripts of approximately 8 kb and 10 kb were detected in thymus, spleen, ovary and testis. In addition, a 1.3 kb transcript was detected in testis.

D. RT-PCR analysis of expression of ERa and ER β in cell lines.

RNA was isolated from a number of human and animal cell lines using RNAzol B (Cinna/Biotecx), cDNA was made using 2.5 microgram of total RNA using the Superscript II kit (BRL) following the manufacturers instructions. A portion of the cDNA was used for specific PCR amplifications of fragments corresponding either to mRNA encoding the ER or to the novel estrogen receptor. (It should be emphasized that the primers used are based on human and rat sequences, whereas some of the cell lines were not rat or human, see legend of Figure 4). Primers used were for ERa: sense 5'-GATGGGCTTACTGACCAACC-3' and antisense 5'-AGATGCTCCATGCCTTTG-3' generating a 548 base pair fragment corresponding to part of the LBD. For ERβ: sense 5'- TTCACCGAGGCCTCCATGATG-3' and antisense 5'-CAGATGTTCCATGCCCTTGTT-3' generating a 565 base pair fragment corresponding to part of the LBD. The PCR samples were analysed on agarose which were blotted onto Nylon membranes. These blots were hybridised with ³²Plabeled PCR fragments generated with the above-mentioned primers on ER α and ER β plasmid DNA using standard experimental procedures (Sambrook et al, 1989).

E. Ligand-dependent transcription activation by the novel estrogen receptor protein.

45

55

Chinese Hamster Ovary (CHO K1) cells were obtained from ATCC (CCL61) and maintained at 37°C in a humidified Cell culture atmosphere (5% CO₂) as a monolayer culture in fenolred-free M505 medium. The latter medium consists of a mixture (1:1) of Dulbecco's Modified Eagle's Medium (DMEM, Gibco 074-200) and Nutrient Medium F12 (Ham's F12, Gibco 074-1700) supplemented with 2.5 mg/ml sodium carbonate (Baker), 55 μ g/ml sodium pyruvate (Fluka), 2.3 μ g/ml β mercaptoethanol (Baker), 1.2 μg/ml ethanolamine (Baker), 360 μg/ml L-glutamine (Merck), 0.45 μg/ml sodium selenite (Fluka), 62.5 μg/ml penicillin (Mycopharm), 62.5 μg/ml streptomycin (Serva), and 5% charcoal-treated bovine calf serum (Hyclone).

Recombinant vectors

The ERß-encoding sequence as presented in SEO ID No. 1 was amplified by PCR using oligonucleotides 5'-

CTTGGATCCATAGCCCTGCTGTGATGAATTACAG-3' (SEQ ID NO:22 underlined is the translation initiation codon) in combination with 5'-GATGGATCCTCACCTCAGGGCCAGGCGTCACTG-3' (SEQ ID NO:23) (underlined is the translation stopcodon, antisense). The resulting BamH1 fragment (approximately 1450 base pairs) were then cloned in the mammalian cell expression vector pNGV1 (Genbank accession No. X99274).

An expression construct encoding the ERβ reading frame as presented in SEQ ID NO:24 was made by replacing a BamH1-Msc1 fragment (nucleotides 1-81 in SEQ ID No. 1) by a BamH1-Msc1 fragment corresponding to nucleotides 77-316 in SEQ ID No. 24. The latter fragment was made by PCR with SEQ ID NO:26 in combination with SEQ ID NO: 28 using the above mentioned 5' RACE fragment.

The reporter vector was based on the rat oxytocin gene regulatory region (position -363/+16 as a HindIII/ Mbol fragment; R.Ivell, and D.Richter, Proc.Natl. Acad.Sci.USA <u>81</u>, 2006-2010, 1984) linked to the firefly luciferase encoding sequence; the regulatory region of the oxytocin gene was shown to possess functional estrogen hormone response elements *in vitro* for both the rat (R. Adan *et al*, Biochem.Biophys.Res.Comm. <u>175</u>, 117-122, 1991) and the human (S. Richard and H.Zingg, J.Biol.Chem. <u>265</u>, 6098-6103, 1990).

Transient transfection

1000

and the same and here

 1×10^5 CHO cells were seeded in 6-wells Nunclon tissue culture plates and DNA was introduced by use of lipofectin (Gibco BRL). Hereto, the DNA (1 μg of both receptor and reporter vector in 250 μL Optimem, Gibco BRL) was mixed with an equal volume of lipofectin reagent (7 μL in 250 μL Optimem, Gibco) and allowed to stand at room temperature for 15 min. After washing the cells twice with serum-free medium (M505) new medium (500 μL Optimem, Gibco) was added to the cells followed by the dropwise addition of the DNA-tipofectin mixture. After incubation for a 5 hour period at 37°C cells were washed twice with fenoIred-free M505 + 5% charcoal-treated bovine calf serum and incubated overnight at 37°C. After 24 hours hormones were added to the medium (10^{-7} mol/L). Cell extracts were made 48 hours posttransfection by the addition of 200 μL lysisbuffer (0.1 M phosphate buffer pH7.8, 0.2% Triton X-100). After incubation for 5 min at 37°C the cell suspension was centrifuged (Eppendorf centrifuge, 5 min) and 20 μL sample was added to 50 μL luciferase assay reagent (Promega). Light emission was measured in a luminometer (Berthold Biolumat) for 10 sec at 562 nm.

Stable transfection of the novel estrogen receptor.

The expression plasmid encoding full-length ER β 1-530 (see above) was stably transfected in CHO K1 cells as previously described (Theunissen *et al.*, J. Biol. Chem. 268, 9035-9040, 1993). Single cell clones that were obtained this way were screened by transient transfection of the reporter plasmid (rat oxytocin-luciferase) as described above. Selected clones were used for a second stable transfection of the rat oxytocin-luciferase reporter plasmid together with the plasmid pDR2A which contains a hygromycine resitance gene for selection. Single cell clones obtained were tested for a response to 17 β -estradiol. Subsequently, a selected single cell clone was used for transactivation studies. Briefly, cells were seeded in 96-wells at (1.6x10⁴ cells per well). After 24 hours different concentrations of hormone were diluted in medium and added to the wells. For antagonistic experiments, 2x10⁻¹⁰ M. 17 β -estradiol was added to each well and different concentrations of antagonists were added. Cells were washed once with PBS after a 24 hour incubation and then lysed by the addition of 40 microliter lysis buffer (see above). Luciferase reagent was added (50 microliter) to each well and light emission was measured using the Topcount (Packard).

Results.

30

40

45

A comparison of the two expression constructs (SEQ ID NO:1 and SEQ ID NO:24) in transient transfections in CHO cells showed identical transactivation in response to a number of agonists and antagonists. CHO cells transiently transfected with ERβ expression vector and a reporter plasmid showed a 3 to 4 fold increase in luciferase activity in response to 17β-estradiol as compared to untreated cells (see Figure 2). A similar transactivation was obtained upon treatment with estriol and estrone. The results indicate not only that the novel ER (ERβ) can bind estrogen hormones but also that the ligand-activated receptor can bind to the estrogen-response elements (EREs) within the rat oxytocin promoter and activate transcription of the luciferase reporter gene. Figure 3 shows that in an independent similar experiment 10⁻⁹ mol/L 17β-estradiol gave an 18-fold stimulation with ERβ and a 7-fold stimulation with ERβ. In addition, the antiestrogen ICI-164384 was shown to be an antagonist for both ERα and ERβ when activated with 17β-estradiol, whereas the antagonist alone had no effect. In this experiment 0.25 μg β-galactosidase vector was co-transfected in order to normalize for differences in transfection efficiency.

Transactivation studies performed on stably transfected ER α and ER β cell lines gave similar absolute luciferase values. The curves for 17 β -estradiol are very similar and show that half-maximal transactivation is reached with lower concentrations of hormone on ER α as compared to ER β (Figure 5). For Org4094 this is also the case however, the

effect observed is much more pronounced. The curves for raloxifen show that the potency of this antagonist to block transactivation on $\mathsf{ER}\alpha$ is greater compared to its potency to block $\mathsf{ER}\beta$ transactivation

Marie Marie

SEQUENCE LISTING

	(1) GENERAL INFORMATION:
10	(i) APPLICANT:
10	(A) NAME: Akzo nobel n.v.
	(B) STREET: Velperweg 76
	(C) CITY: Arnhem
15	(E) COUNTRY: The Netherlands
	(F) POSTAL CODE (ZIP): 6824 BM
	(G) TELEPHONE: 0412-666379
	(H) TELEFAX: 0412-650592
20	(I) TELEX: 37503 akpha nl
	(ii) TITLE OF INVENTION: Novel estrogen receptor
25	(iii) NUMBER OF SEQUENCES: 28
	(iv) COMPUTER READABLE FORM:
30	(A) MEDIUM TYPE: Floppy disk
	(B) COMPUTER: IBM PC compatible
	(C) OPERATING SYSTEM: PC-DOS/MS-DOS
	(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
35	
	(2) INFORMATION FOR SEQ ID NO: 1:
40	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 1434 base pairs
	(B) TYPE: nucleic acid
45	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear
	(ii) MOLECULE TYPE: cDNA
50	
	(xi) SECUENCE DESCRIPTION. SEC ID NO. 1.

	ATGANTTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC	80
5	ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC	120
	CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA	180
10	GAACACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC	240
15	GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC	300
	GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTT	360
20	AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC	420
	GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA	480
25	ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA	540
	AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC	600
30	CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG	660
35	GAGGCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC	720
33	ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC	780
40	AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC	840
	TGTTGGATGG AGGTGTTAAT GATGGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG	900
45	CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT	960
	CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAA	1020 •
50	CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG	1080
	GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG	1140

55

÷. .

and the second s

was a maganitation to the

Angelow Commence

But the state of the state of the state of 1932年在中国各中的中国

55

	ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG	1200
5	CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGTAA CAAGGGCATG	1260
	GAACATCTGC TCAACATGAA GTGCAAAAAT GTGGTCCCAG TGTATGACCT GCTGCTGGAG	1320
10	ATGCTGAATG CCCACGTGCT TCGCGGGTGC AAGTCCTCCA TCACGGGGTC CGAGTGCAGC	1380
	CCGGCAGAGG ACAGTAAAAG CAAAGAGGGC TCCCAGAACC CACAGTCTCA GTGA	1434
15	(2) INFORMATION FOR SEQ ID NO: 2:	
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1251 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
25	(ii) MOLECULE TYPE: cDNA	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
35	ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC	60
	ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC	120
40	CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA	180
	GAACACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC	240
45	GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC	300
50	GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT	* 360
50	AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC	420
	GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA	480

A Secretary of Secretary of the secretar

5

	ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA	540
	ACTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC	600
	CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG	660
0	CACCCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC	720
	ATCATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC	780
15	AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC	840
20	TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG	900
20	CTCATCTTIG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT	960
25	CTGGAAAICT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AMAACTCCAA	1020
	CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG	1080
30	GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG	1140
	ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG	1200
35	CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTG A	1251
40	(2) INFORMATION FOR SEQ ID NO: 3:	
4	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 66 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	•

(ii) MOLECULE TYPE: peptide

55

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
5	Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp
10	Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His
15	Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn 35 40 45
	Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val 50 55 60
20	Gly Met 65
25	(2) INFORMATION FOR SEQ ID NO: 4:
30	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 233 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single
35	(D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
45	Leu Val Leu Thr Leu Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser 1 5 10 15
50	Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser Met Met Net Ser Leu Thr 20 25 30
	Lys Leu Ala Asp Lys Glu Leu Val His Met Ile Ser Trp Ala Lys Lys 35 40 45

at visitaret er est Timoret er en er Timoret er en er

A property of the second of th

55

	Ile	Pro 50	Gly	Phe	Val	Glu	Leu 55	Ser	Leu	Phe	Asp	Gln 60	Val	Arg	Leu	Leu
5	Glu 65	Ser	Cys	Trp	Met	Glu 70	Val	Leu	Met	Met	Gly 75	Leu	Met	Trp	Arg	Ser 80
10	Ile	qeA	His	Pro	Gly 85	Lys	Leu	Ile	Phe	Ala 90	Pro	Asp	Leu	Val	Leu 95	Asp
15	Arg	Asp	Glu	Gly 100	Lys	Cys	Val	Glu	Gly 105	Ile	Leu	Glu	Ile	Phe 110	Asp	Met
20	Leu	Leu	Ala 115		Thr	Ser	Arg	Phe 120		Glu	Leu	Lys	Leu 125	Gln	His	Lys
25	Glu	Туг 130		Cys	Val	Lys	Ala 135		Ile	Leu	Leu	Asn 140	Ser	Ser	Met	Tyr
	Pro	Let	ı Val	Thi	Ala	The	Glr	Asp	Ala	Asp	Ser	Ser	Arg	Lys	Leu	Ala
30	145					150					155					160
30	Hi	s Le	u Let	ı Ası	n Ala 16		L Thi	r As _l	p Ala	170		. Trį	Val	. Ile	175	Lys
35	Se	r Gl	y Il	e Se 18		r Gl	n Gl	n Gl	n Se 18		t Ar	g Le	a Ala	19	Let	ı Leu
40	Ме	t Le	u Le		r Hi	s Va	l Ar	g Hi 20		a Se	r As	n Ly	s Gl 20	у Ме 5	t Gl	u His
45	L€		eu As 10	sn Me	et Ly	ys Cj		γ s A s 15	sn Vá	al Vē	al Pr	o Va 22	1 Ty 10	r As	p Le	u Leu
	L	eu G	lu M	et L	eu A	sn A	la H	is V	al L	eu						
50		25					30									
(2) IN	FORM	OITA	n fo	R SE	Q ID	NO:	5:								

(i) SEQUENCE CHARACTERISTICS:

And the second

**.

.

	(A) LENGTH: 477 amino acids
	(B) TYPE: amino acid
5	
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: unknown
10	(ii) MOLECULE TYPE: protein
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
20	Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Pro 1 10 15
25	Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro Thr Pro Gly His
30	Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His Leu Tyr Ala Glu 35 40 45
	Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu Glu His Thr Leu 50 55 60
35	Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser Gly Asn Arg Cys 65 70 75 80
40	Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His Phe Cys 85 90 95
45	Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser 100 105 110
50	Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn 115 120 125
	Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg
	130 135 140
55	T30

The second secon

	Ten Arg Leu Arg Lys Cys Tyr Glu Val Gly
	Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly
	145
5	Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys Gly Tyr Arg Leu Val
	165
	Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala
10	
	180
	Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Leu Asp
	Lys Arg Ser Gly Gly His Ala Pro Arg 1205
15	195
	Tou Lou Glu Ala Glu Pro
	Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro
	210 215
20	
	Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser 235 240
	Pro His Val Leu 220 235
	225
25	Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met 250 255
	Met Met Ser Leu Thr Lys Deu 1250
	245
	Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
30	He Ser Trp Ala Lys Lys He Pro Gly Fne val 270
	260 265
	-1 Val Tau Met Net
35	Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
30	280
	Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
40	245
	290
	Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile 315 320
	Pro Asp Leu Val Leu Asp Arg Asp 315
45	305
	Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
	Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Inc 335
	325
50	as - Mat Tle Tell
	Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
	345 345
•	330

	Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp 355 360 365
5	Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu 370 375 380
10	Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Het 395 390 395 400
15	Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Ser 405 410 415
20	Asn Lys Gly Met Glu His Leu Leu Asn Met Lys Cys Lys Asn Val Val 420 425 430
	Pro Val Tyr Asp Leu Leu Glu Met Leu Asn Ala His Val Leu Arg 435 440 445
25	Gly Cys Lys Ser Ser Ile Thr Gly Ser Glu Cys Ser Pro Ala Glu Asp 450 455 460
30	Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln Ser Gln 465 470 475
35	2) INFORMATION FOR SEQ ID NO: 6:
	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 416 amino acids
-1 0	(B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown
45	(ii) MOLECULE TYPE: protein
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
5 <i>5</i>	Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Pro

	10	15
	1	
	a san Val Leu Trp Pro	Thr Pro Gly His
5	Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro	
10	Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His 35	
	Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu 50	
15	and I've Val Sel	Gly Asn Arg Cys
	Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Sei 70	
20	Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg As	
25	Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Ty	
	the Dhe Lys Arg Ser I	le Gln Gly His Asn
30	Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser I 115	125
		lle Asp Lys Asn Arg
	Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr	140
35	130	
35	Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys	Cys Tyr Glu Val Gly 160
	150	5 Val
40	Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys 165	
	ar Iou Wil	CVS Ala Gly Lys Ala
45	Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His	
	Lys Arg Ser Gly Gly His Ala Pro Arg Val Ar	g Glu Leu Leu Leu Asp
50	195	
	Ala Leu Ser Pro Glu Gln Leu Val Leu Thr L	eu Leu Glu Ala Glu Pro
55	ALA DEL DO-	

	220
	215
	210
	Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser 230 235
	Win Val Leu Ile Ser Arg Pro Ser Ala 125
	230 230
5	
	asp Lys Glu Leu Val His Mee
	yet Ser Leu Thr Lys Leu Ala Asp 255
	225 Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met 255 245
.0	
10	Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe 270 265
	Sor Tro Ala Lys Lys Ile Pio Gly 270
	265
	200
15	Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met 285
	ach Gln Val Arg Leu Leu Glu Ser of
	213
	Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala 295
20	Gly Leu Met Trp Arg Ser 11e 7mp 300
	- **
	250 Glu Gly Ile
	Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile 320 310
25	Pro Asp Leu Val Leu Asp 123 315
25	
	Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu 335
	335
	Leu Glu Ile Phe Asp 1113
30	343
	Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu 350
	rau Gln His Lys Glu Tyr Leu Cys var 350
	Leu Lys Leu 345
	340
35	Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp 365
	365
	Leu Asir 505
	355
40	Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu 375
40	ser Ser Arg Lys Leu Ala His Leu 380
	370
	400
45	Val Trp Val Ile Ala Bys 555 395
	385 Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Arg 415
	A15
	Arg Leu Ala Ash Leu 410
50	405

		40	
		EP 0 798 378 A2	
25450			
(A)		FORMATION FOR SEQ ID NO: 7:	
	(2) IN	IEO Kasur zan	
with the state of		(i) SEQUENCE CHARACTERISTICS:	
	5		
		nucleic ad-	
		STRANDEDNESS:	
	10	(D) TOPOLOGY: unknown	
	,		
		(ii) MOLECULE TYPE: CDNA	
		(11)	
	15		
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
		(xi) SEQUENCE DESCRIPTION:	29
		and Mark Ville	
	20	EIGAYGARG CWTCIGGITG YCAYTAYGG	
	60	TD NO: 8:	
	1	2) INFORMATION FOR SEQ ID NO: 8:	
	25	TO THE STATE OF TH	
		(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 29 base pairs	
		(A) LENGTH: 25 Date (B) TYPE: nucleic acid	
		(B) TYPE: nucleon (C) STRANDEDNESS: single	
	30	(C) STRANDELINGS: linear (D) TOPOLOGY: linear	
		(D) TOPOLOGI: 22	
		TYPE: CINA	
	35	(ii) MOLECULE TYPE: CUNA	
	55		
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:	29
	40	(xi) SEQUENCE	2
	•	AAGCCTGGSA YICKYTTIGC CCAIYTIAT	
		AAGCCTGGSA	
		(2) INFORMATION FOR SEQ ID NO: 9:	•
	. 45	(2) INFORMATION	
		(i) SEQUENCE CHARACTERISTICS:	
	50	and ell about	
		CERANDEDNESS: 32113	
		(D) TOPOLOGY: linear	
		•	
	55		

	(ii) MOLECULE TYPE: cDNA	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:	22
0 TG	TTACGAAG TGGGAATGGT GA	
) INFORMATION FOR SEQ ID NO: 10:	
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
25	(11) MOLECULE TYPE: CDNA	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:	24
	TTGACACCAG ACCAACTGGT AATG	
35	(2) INFORMATION FOR SEQ ID NO: 11:	
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
45	(ii) MOLECULE TYPE: cDNA	•
<i>.</i>	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:	
	55	

GGTG	GCGACG ACTCCTGGAG CCCG	
5 (2)	INFORMATION FOR SEQ ID NO: 12:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
15	(ii) MOLECULE TYPE: cDNA	
20	(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 12:	22
25	GTACACTGAT TTGTAGCTGG AC (2) INFORMATION FOR SEQ ID NO: 13:	
30	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	
35	(D) TOPOLOGY: linear (ii) MOLECULE TYPE: CDNA	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:	20
45	CCATGATGAT GICCCTGACC	:
50	(2) INFORMATION FOR SEQ ID NO: 14: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid	

And the second of the second

	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
5	(ii) MOLECULE TYPE: cDNA	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:	
15	TCGCATGCCT GACGTGGGAC	20
	(2) INFORMATION FOR SEQ ID NO: 15:	
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 base pairs	
25	(B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
30	(ii) MOLECULE TYPE: cDNA	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:	24
	GGCSTCCAGC ATCTCCAGSA RCAG	
40	(2) INFORMATION FOR SEQ ID NO: 16:	
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	
5	(D) TOPOLOGY: linear	•
	(ii) MOLECULE TYPE: cDNA	

A Marie Comment

The color of the section of And the second second

	SEO ID NO: 16:	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:	20
		20
5	GCAAGCTGGC TCACTTGCTG	
	(2) INFORMATION FOR SEQ ID NO: 17:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
,,,	(D) TOPOLOGY: linear	
	(11) MOLECULE TYPE: CDNA	
20		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:	
25	(XI) SDE-	20
	TCTTGTTCTG GACAGGGATG	
30	(2) INFORMATION FOR SEQ ID NO: 18:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 20 base pairs	
35	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
40	(ii) MOLECULE TYPE: cDNA	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:	20
	GCATGGAACA TCTGCTCAAC	
	(2) INFORMATION FOR SEQ ID NO: 19:	
	(i) SEQUENCE CHARACTERISTICS:	

.

	(A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: cDNA	
· · · · · ·	10	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:	
	AGCAAGTTCA GCCTGTTAAG T	
	20 (2) INFORMATION FOR SEQ ID NO: 20:	
	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1257 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: CDNA	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:	60
	(xi) SEQUENCE DESCRIPTION: SEQ TO ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT AAGATCGCTA	120
	ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGGTCCCT GGTGTGAAGC AAGATCGCTA CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGG GAACCGTTGC	240
	GAACACACCT TACCTGTAAA CAGAGAGACA CIC	300 360
	GCCAGCCCTG TTACTGGTCC AGGTTCARAG AGGGTT GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT	300
	GATTACGCAT CGGGG	

AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC	420
ATTTGTCCAG CTACAAATCA GIGILA	
TTCAAGGACA TAATGATTAT ATTT	480
AAAAGAAGCA TTCAAGGACA TAATGATTA GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA	
TEGERAGIC TGCCGACTIC	
CATAAAAACC GGCGCAAGAG CIGGCACAGAGA	540
GATAAAAACC GGCGCAAGAG CTGCCACGCGCAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGGTACC GCCTTGTGCG GAGACAGAGA	
CTCCCTCCCG GAGAGAGA 19100	600
ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA GAAGTGGCGG CCACGCGCCC AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCG CCACGCCCCC	
TOTAL CONTROL OF THE PROPERTY	
O ACCCTCCTG	660
AGTGCCGACG AGCAGCTGCA CTGTGCCGCGAGC AGCTAGTGCT CACCCTCCTG CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG	
ACCTGCTGCT GGACGCCCTG AGGGGG	720
CGAGTGCGGG AGGICTICAC CGAGGCCTCC	
TOTAL CACC COCCCAGIO	-00
SACCOTGAGO CGCCCCATGT GCTGAG	780
GAGGCTGAGC CGCCCCATGT GCTGATCHE ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC	
COUTGACCAA GTTGGCCGAL ANGULA	840
ATGATGATGT CCCTGACCAA GTTGGCCCCCC 20 AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CCCCGGCAAG	
20 CTGTTCGGC CTGTTCGGC	- • •
SACREGATTC CCGGCTTTGT GGGCAAG	900
AMOTOGCOCT CAATTGACCA COUR	
AAGAAGATTC CCGGCTTTGT GGARCICATO AAGAAGATTC CCGGGCTTGT GGARCICATO TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG	960
TGTTGGATGG AGGAGGAATA	
AGGGATTANUS	
TGTTGGATGG AGGTGTTAAT GATGGGGGAT 25 CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT CATCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG TTCGAGAGTT AAAACTCCA	A 1020
CTCATCTTTG CTCCAGATCT TGTTCTGGC 30 CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAG ATTCCAGTAT GTACCCTCT	
TOTAL CATGOT COTGGCAACT ACTION	1080
CTGGAAATCT TTGACATGCT CCTGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCT CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCT	G
CCCCATG ATCCTGCTCA	
CARAGAAT ATCTCTGTGT CAACGCCG	TG 1140
GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCG	
CCCACCATGC TGACAGCAGC CGGGAAGC	1200
35 GTCACAGCGA CCCAGGCA GCAATCC	ATG
CATTGCCAAG AGCGGCATCT	
GTCACAGCGA CCCAGGATGC TGACAGCTAG GTCACAGCGA CCCAGGATGC TGACAGCTAGCAAG AGCGGCATCT CCTCCCAGCA GCAATCC ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCC	1257
ACCOMPANY CACCTCAGGC ATGCGAGGTC 1000	
40 ACCITICATE GCTCCTGTCC CACCO	
ACCGATGCTT TGGTTTGGGT GATTOTA 40 CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTC TGCCTGF	
(2) INFORMATION FOR SEQ ID NO: 21:	
INFORMATION FOR SEQ 12	
(2)	
(i) SEQUENCE CHARACTERISTICS:	•
(i) SEQUENCE CHARGET And amino acids (A) LENGTH: 418 amino acids	
(A) LENGTH. acid	
(B) TYPE: amino acid	
50 STRANDEDNESS: SILVE	
(D) TOPOLOGY: linear	
(N) 10222	
55 (ii) MOLECULE TYPE: protein	
55 (11)	

			- *D NO:	21:	
137 THE WAY SHEET !		DESCRIPT	ION: SEO ID NO.		
	5	(xi) SEQUENCE DESCRIPT		Gl1	Gly Gly Pro
			ar Nen Val	Thr Asn Leu GI	15
* *		Tur Ser Ile I	Pro Ser Asi	10	13
		Met Ash 191 5		10	
		Gly Arg Gln Thr Thr		-1.	- Pro Gly His
	10	-	Val	Leu Trp Pro Tr	ir fic or:
ű.		Thr Thr	Ser Pro Ash var		30
-		Gly Arg Gin Int	25		
		20			ala Glu
		20 Leu Ser Pro Leu Val	a1 =	Leu Ser His L	eu Tyr Ala
	15	Val	Val His Arg Gin	4	15
		Leu Ser Pro Leu van	40		
		35			mu TAN
		35 Pro Gln Lys Ser Pr		ger Leu	Glu His Thr Dea
			TED CVS Glu AL	a Arg Jer -	
	20	neo Gln Lys Ser Pr	O 115 -1	60	
	20	PIG GEN	55		
		50 Pro Val Asn Arg G			Gly Asn Arg Cys
			Tou Lus A	rg Lys Val Ser	80
		Wal Asp Arg G	lu Thr Lau 52	75	
	25	Pro Val	70		
• • •	25	65 Ala Ser Pro Val			ale wis Phe Cys
			m1	Ser Lys Arg Asi	Ala III
		Val	Lpr CJA bro gra .	202 -2	95
		Ala Ser Pro Val	A.E.	90	
	_		85		yr Gly Val Trp Ser 110
	30			TUT HIS T	Ir Gly Val IIP
			Asp Tyr Ala Ser	GIA 135	110
		Ala Val Cys Ser	Yah	105	
		100			
				ger I	tle Gln Gly His Asn 125
	35		Ala Phe Phe	e Lys Arg Ser	125
•		Cur Glu Gly Cys	Lys Alu	Λ	123
		Cys 02.1	12	•	
* .		113			Ile Asp Lys Asn Arg 140
			Thr At	sn Gln Cys Thr	116
	40	mur tle Cy	s Pro Ala III	-	140
		Asp Ty2	135		
· n		130			Cya Tyr Glu Val Gly
•				Leu Arg Lys	160
		gar C	vs Gln Ala Cys /	155	5
	45	Arg Lys Ser S	150		
		145			ard Leu Val
		110		Clu Ard Cy	s Gly Tyr Ary 200
			cua Gly Ser Arg	Arg Gru	s Gly Tyr Arg Leu Val 175
		Met Val Lys	Cya des	170	
	50		165		
				· ·	is Cys Ala Gly Lys Ala
			car Ala ASI	Glu Gln Leu n	is Cys Ala Gly Lys Ala
		Arg Arg Gln	Arg Ser /		
	*	Vrd >			
	55				
	50				

			190	
	180	185		
	Lys Arg Ser Gly Gly His Ala		Glu Leu Leu Leu A	sp.
	sor Gly Gly His Ala	Pro Arg Val Arg	205	
5	Lys Arg Ser 321	200	200	
5	195		- ala Glu F	ro
	Ala Leu Ser Pro Glu Gln Leu	Val Leu Thr Leu	Leu Glu Ala Gra	
	Ala Leu Ser Pro Glu Gin 1655	•	220	
10	Pro His Val Leu Ile Ser Arc	our ser Ala Pro	Phe Thr Glu Ala	Ser
	Pro His Val Leu Ile Ser Are	235	ò	240
	225 Met Met Met Ser Leu Thr Ly	T.V	a Glu Leu Val His	Met
15	war Ser Leu Thr Ly	s Leu Ala Asp Dy	255	
	Ile Ser Trp Ala Lys Lys I		Ser Leu	Ph•
	TVE LVS I	le Pro Gly Phe Va	270	
20	Ile Ser Trp Ala Lys 270	265	270	
	Asp Gla Val Arg Lau Lou	clu Ser Cys Trp N	est Glu Val Leu Par	.
	Asp Gla Val Arg Lau Lau	280	285	
25	A7E			
	Gly Leu Met Trp Arg Ser	Tie Pro	Gly Lys Leu Ile Ph	e VIS
	Cly Lau Met Trp Arg Ser	Ile Asp nis	300	
30	230		Twe Cvs Val Glu G	ly Ile
	290 Pro Asp Leu Val Leu Asp	Arg Asp Glu Gly	DIS CIT	320
	32.			
	305 Leu Glu Ile Phe Asp Me		ara Phe	Arg Glu
35	all hen Me	t Leu Leu Ala Thr	The Ser Ary	335
	Leu Glu Ile Phe Amp	330)	
	323			Tle Leu
	Leu Lys Leu Gln His L	Glu Tyr Leu Cy	s Val Lys Ala Met	110 23
40	Leu Lys Leu Gln His L	345	350	
	Leu Asn Ser Ser Met 1	Val T	hr Ala Thr Gln Asp	Ala Asp
	Teu Asn Ser Ser Met 1	lyr pro Led var	365	;
45				
		_ ,	non his Val Thr As	p Ala Leu
	355 Ser Ser Arg Lys Leu	Ala His Leu Leu	380	
50	370		- cin Gin G	ln Ser Met
	370 Val Trp Val Ile Ala	Lys Ser Gly Ile	Ser Ser Gin Gin G	
	Val Trp Val 11e Ala	• •		
55				

	390	400
	385	ard Ard
5	Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His 405	Val Arg His Ala Aug 415
10	Ser Ala	
	(2) INFORMATION FOR SEQ ID NO: 22:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs	
	(R) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
20	(D) TOPOLOGY: linear	
25	(ii) MOLECULE TYPE: cDNA	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:	34
	CTTGGATCCA TAGCCCTGCT GTGATGAATT ACAG	
35	(2) INFORMATION FOR SEQ ID NO: 23:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 33 base pairs	
40	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
45	(ii) MOLECULE TYPE: cDNA	:
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:	23:
*	GATGGATCCT CACCTCAGGG CCAGGCGTCA CTG	
55		

(2) INFORMATION FOR SEQ ID NO: 24:

The state of the s

and the make a section of the last the

The same with the same of the same

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1898 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

15

25

30

35

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CACGAATCTT TGAGAACATT ATAATGACCT TTGTGCCTCT TCTTGCAAGG TGTTTTCTCA 60 GCTGTTATCT CAAGACATGG ATATAAAAAA CTCACCATCT AGCCTTAATT CTCCTTCCTC 120 CTACAACTGC AGTCAATCCA TCTTACCCCT GGAGCACGGC TCCATATACA TACCTTCCTC 180 CTATGTAGAC AGCCACCATG AATATCCAGC CATGACATTC TATAGCCCTG CTGTGATGAA 240 TTACAGCATT CCCAGCAATG TCACTAACTT GGAAGGTGGG CCTGGTCGGC AGACCACAAG 300 CCCAAATGTG TTGTGGCCAA CACCTGGGCA CCTTTCTCCT TTAGTGGTCC ATCGCCAGTT 360 ATCACATCTG TATGCGGAAC CTCAAAAGAG TCCCTGGTGT GAAGCAAGAT CGCTAGAACA 420 CACCTTACCT GTAAACAGAG AGACACTGAA AAGGAAGGTT AGTGGGAACC GTTGCGCCAG 480 CCCTGTTACT GGTCCAGGTT CAAAGAGGGA TGCTCACTTC TGCGCTGTCT GCAGCGATTA 540 600 CGCATCGGGA TATCACTATG GAGTCTGGTC GTGTGAAGGA TGTAAGGCCT TTTTTAAAAG AAGCATTCAA GGACATAATG ATTATATTTG TCCAGCTACA AATCAGTGTA CAATCGATAA 660 AAACCGGCGC AAGAGCTGCC AGGCCTGCCG ACTTCGGAAG TGTTACGAAG TGGGAATGGT 720

GAAGTGTGGC	TCCCGGAGAG	AGAGA1G1GG	GIACCGCCII	010coananc	AGAGAAG1 GC	, 00
CGACGAGCAG	CTGCACTGTG	CCGGCAAGGC	CAAGAGAAGT	GGCGGCCACG	CGCCCCGAGT	840
GCGGGAGCTG	CTGCTGGACG	CCCTGAGCCC	CGAGCAGCTA	GTGCTCACCC	TCCTGGAGGC	900
TGAGCCGCCC	CATGTGCTGA	TCAGCCGCCC	CAGTGCGCCC	TTCACCGAGG	CCTCCATGAT	960
GATGTCCCTG	ACCAAGTTGG	CCGACAAGGA	GTTGGTACAC	ATGATCAGCT	GGGCCAAGAA	1020
GATTCCCGGC	TTTGTGGAGC	TCAGCCTGTT	CGACCAAGTG	CGGCTCTTGG	AGAGCTGTTG	1080
GATGGAGGTG	TTAATGATGG	GGCTGATGTG	GCGCTCAATT	GACCACCCCG	GCAAGCTCAT	1140
CTTTGCTCCA	GATCTTGTTC	TGGACAGGGA	TGAGGGGAAA	TGCGTAGAAG	GAATTCTGGA	1200
AATCTTTGAC	ATGCTCCTGG	CAACTACTTC	AAGGTTTCGA	GAGTTAAAAC	TCCAACACAA	1260
AGAATATCTC	TGTGTCAAGG	CCATGATCCT	GCTCAATTCC	AGTATGTACC	CTCTGGTCAC	1320
AGCGACCCAG	GATGCTGACA	GCAGCCGGAA	GCTGGCTCAC	TTGCTGAACG	CCGTGACCGA	1380
TGCTTTGGTT	TGGGTGATTG	CCAAGAGCGG	CATCTCCTCC	CAGCAGCAAT	CCATGCGCCT	1440
GGCTAACCTC	CTGATGCTCC	TGTCCCACGT	CAGGCATGCG	AGTAACAAGG	GCATGGAACA	1500
TCTGCTCAAC	ATGAAGTGCA	AAAATGTGGT	CCCAGTGTAT	GACCTGCTGC	TGGAGATGCT	1560
GAATGCCCAC	GTGCTTCGCG	GGTGCAAGTC	CTCCATCACG	GGGTCCGAGT	GCAGCCCGGC	1620
AGAGGACAGT	AAAAGCAAAG	AGGGCTCCCA	GAACCCACAG	TCTCAGTGAC	GCCTGGCCCT	1680
GAGGTGAACT	GGCCCACAGA	GGTCACAAGC	TGAAGCGTGA	ACTCCAGTGT	GTCAGGAGCC	1740 \$
TGGGCTTCAT	CTTTCTGCTG	TGTGGTCCCT	CATTTGGTGA	TGGCAGGCTT	GGTCATGTAC	1800
01.maammaaa	mcca ccmmcc	CAACTCTCAG	CACTCCCTCT	GAGGAAGCCA	тасттесст	186

TGTTAGCAGA GGGACATTTG AATCGAGCGT TTCCACAC

5	(2)	INFOR	MATIC	ON FO	OR SE	Į II	NO:	25:	:								
		(i)	SPOU	ENCE	CHAF	LACT!	ZRIST	rics	:								
		(1)			GTH:												
10					E: au												
10					ANDE				e								
					OLOG												
			,-,														
15		(ii)	MOLE	CULE	TYP	E: p	eptic	de									
20																	
20																	
		(xi)	SEQU	ence	DES	CRIP	TION	: SE	Q ID	NO:	25:						
									_	_		•	g	D	9	gar	Tur
25		Met	Asp	Ile			Ser	Pro	Ser	Ser	Leu 10	Asn	ser	Pro		3 6 1 15	.yr
		1				5					10						
			_		41	o	T1 =	7 411	Pro	Leu	G) u	His	Gl v	Ser	Ile	Tyr	Ile
30		Asn	Cys	ser	20	961	TIE	Den	110	25	0		,		30	•	
					20												
		Pro	Ser	Ser	Tvr	Val	Asp	Ser	His	His	Glu	Tyr	Pro	Ala	Met	Thr	Phe
		1.0		35	-1-		•		40					45			
35																	
		Tyr	Ser	Pro	Ala	Val	Met	Asn	Tyr	Ser	Ile	Pro	Ser	Asn	Val	Thr	Asn
			50					55					60				
40																	
		Leu	Glu	Gly	Gly	Pro	Gly	Arg	Gln	Thr	Thr	Ser	Pro	Asn	Val	Leu	Trp
		65					70					75					80
45													•••			T	602
40		Pro	Thr	Pro	Gly		Leu	Ser	Pro	Leu		Val	HIS	Arg	GIN	95	361
						85					90					,,	
				_		63	D	61 -		Car	. Dro	Trn	Cvs	Gli	ı Ala	Ara	Ser
50		His	Leu	Туг			Pro	GTU	гъ	105		, 1-P	-,-	-	110		,
					100	1				10.	•						
		T	. 61.	, Hi-	. The	Lev	Pro	Val	Asr	Arc	g Glu	ı Thr	Leu	Ly:	. Arç	, Lys	Val
55		Der	. 610	. 11+3							-						

and proceedings on the comment of th

5	Ser	Gly 130	As n	Arg	Суз	Ala	Ser 135	Pro	Val	Thr (Pro 140	Gly	Ser	Lys .	Arg
10	Asp 145	Ala	His	Phe	Cys	Ala 150	Val	Сув	Ser		Ty r 155	Ala	Ser	Gly		ні s 160
	Tyr	Gly	Val	Trp	Ser 165	Cys	Glu	Gly	Сув	Lys 170	Ala	Phe	Phe	Lys	Arg 175	Ser
15	Ile	Gln	Gly	His 180	Asn	As p	Tyr	Ile	Cys 185	Pro	Ala	Thr	Asn	Gln 190	Суз	Thr
20	Ile	Asp	Lys 195	Asn	Arg	Arg	Lys	Ser 200	Cys	Gln	Ala	Cys	Arg 205	Leu	Arg	Lys
25	Cys	Tyr 210	Glu	Val	Gly	Met	Val 215	Lys	Cys	Gly	Ser	Arg 220	Arg	Glu	Arg	Cys
30	Gly 225		Arg	Leu	Val	Arg 230	Arg	Gln	Arg	Ser	Ala 235	Asp	Glu	Gln	Leu	His 240
30	Суз	Ala	Gly	Lys	Ala 245		Arg	Ser	Gly	Gly 250		Ala	Pro	Arg	Val 255	Arg
35	Glu	ı Lei	ı Leu	Leu 260		Ala	Lev	. Ser	Pro 265		Gln	Leu	Val	Leu 270		Leu
40	Let	ı Gl	u Ala 275		ı Pro	Pro	Hi:	280		ı Ile	Ser	Arg	285		: Ala	Pro
45	Ph	e Th 29		u Ale	a Se.	r Me	29		t Se	r Let	1 Thi	300		ı Ala	As _i	Lys
50	G1 30		u Va	l Hi	s Me	t Il 31		r Tr	p Al	a Ly:	5 Ly		e Pr	o Gl	y Ph	e Val
50	G1	u Le	eu Se	r Le	u Ph	e As	p Gl	n Va	l Ar	g Le	u Le	u Gl	u Se	r Cy	s Tr	p Met

EP 0 798 378 A2

330

325

335

erweigen einer gewerkelten. Von Weigertreiten eine der

The second secon

5	Glu	Val	Leu	Met	Met	Gly	Leu	Met	Trp	Arg	Ser	Ile	qeA	His	Pro	Gly
				340					345					350		
	Lys	Leu	Ile	Phe	Ala	Pro	Asp	Leu	Val	Leu	Asp	Arg	Asp	Glu	Gly	Lys
0			355					360					365			
	Cys	Val	Glu	Gly	Ile	Leu	Glu	Ile	Phe	Asp	Met	Leu	Leu	Ala	Thr	Thr
5		370					375					380				
•											_					.
		Arg	Phe	Arg	Glu	•	Lys	Leu	Gln	His		Glu	Tyr	Leu	Cys	
	385					390					395					400
20	•	Ala		7 1.	7	7		c		Wat	W	Dra	Tou	Va1	ሞኤ ~	X1 a
	гÀ2	AT.	met	116	405	ren	ABII	Set	3e1	410	ığı	PLO	Leu	Val	415	Λια
					403					410					413	
25	Thr	Gln	Asp	Ala	Asp	Ser	Ser	Arσ	Lvs	Leu	Ala	His	Leu	Leu	Asn	Ala
				420					425					430		
	Val	Thr	Asp	Ala	Leu	Val	Trp	Val	Ile	Ala	Lys	Ser	Gly	Ile	Ser	Ser
30			435					440					445			
	Gln	Gln	Gln	Ser	Met	Arg	Leu	Ala	Asn	Leu	Leu	Met	Leu	Leu	Ser	His
35		450					455					460				
		Arg	His	Ala	Ser		Lys	Gly	Met	Glu		Leu	Leu	Asn	Met	_
	465					470					475					480
40	_	•		17. 1	17.1		17- 1		•	•	*	7	~1	W-+	T	N a m
	Суз	Lys	Asn	VAI	Val 485	Pro	Val	Tyr	Asp	190	Leu	ьеu	GIU	Ret	495	ASII
					403					1,50					455	
45	Ala	His	Val	Leu	Arg	Glv	Cvs	Lvs	Ser	Ser	Ile	Thr	Glv	Ser	Glu	Cys
	,		Vu2	500	•	1	-1-	-,-	505					510		•
									_							
50	Ser	Pro	Ala	Glu	Asp	Ser	Lys	Ser	Lys	Glu	Gly	Ser	Gln	Asn	Pro	Gln
50			515					520					525			
	_	-1														

530

A SAMPLE SAMPLE SAMPLE

A STATE OF THE STA

5	(2) INFORMATION FOR SEQ ID NO: 26:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 30 base pairs	
10	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
15		
	(ii) MOLECULE TYPE: other nucleic acid	
20		
	(ad) analysis programmer and to up of	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:	
25	GTGCGGATCC TCTCAAGACA TGGATATAAA	30
	Olocomics Islandian Islandian	3(
	(2) INFORMATION FOR SEQ ID NO: 27:	
30		
	(1) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 25 base pairs	
	(B) TYPE: nucleic acid	
35	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
40	(ii) MOLECULE TYPE: other nucleic acid	
45		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:	
	ACTANCACCC CTCCCCCAAC CCTTC	2!
50	AGTAACAGGG CTGGCGCAAC GGTTC	2:
	(2) INFORMATION FOR SEQ ID NO: 28:	
55	(i) SEQUENCE CHARACTERISTICS:	

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

15

20

10

and the second section in the

Contract to the first for the

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

ACTGGCGATG GACCACTAAA GG

22

25 Claims

- Isolated DNA encoding a protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said protein exhibits at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and the amino acid sequence of said ligand-binding domain of said protein exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4.
- Isolated DNA according to claims 1, characterized in that the amino acid sequence of said DNA-binding domain
 of said protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with
 the amino acid sequence shown in SEQ ID NO:3.

35

30

- Isolated DNA according to claims 1 or 2, characterized in that the amino acid sequence of said ligand-binding domain of said protein exhibits at least 75%, preferably 80%, more preferably 90%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO:4.
- Isolated DNA according to claims 1 to 3, said DNA encoding a protein comprising the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.
 - Isolated DNA according to claims 1 to 4, characterized in that said DNA comprises the nucleic acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20 or SEQ ID NO:24.

- 6. A recombinant expression vector comprising the DNA according to any of the claims 1 to 5.
- 7. A cell transfected with DNA according to claims 1 to 5 or an expression vector according to claim 6.
- 8. A cell according to claim 7 which is a stable transfected cell line which expresses the steroid receptor protein according to any of the claims 9 to 11.
 - 9. Protein encoded by DNA according to claims 1 to 5 or an expression vector according to claim 6.
- Protein according to claim 9, said protein comprising the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.
 - 11. Chimeric protein having an N-terminal domain, a DNA-binding domain, and a ligand-binding domain,

EP 0 798 378 A2

characterized in that at least one of said domains of said chimeric protein originates from a protein according to claims 9 or 10, and at least one of the other domains of said chimeric protein originates from another receptor protein from the nuclear receptor superfamily, provided that the DNA-binding domain and the ligand-binding domain of said chimeric protein originates from different proteins.

12. DNA encoding a protein according to claim 11.

Commence of the

5

10

15

20

25

30

35

40

50

55

estance for Thispen Cyni

- 13. Use of a DNA according to claims 1 to 5 or 12, an expression vector according to claim 6, a cell according to claim 7 or 8 or a protein according to claim 9 to 11 in a screening assay for identification of new drugs.
- 14. A method for identifying functional ligands for the protein according to claims 9 to 11, said method comprising the steps of
- a) introducing into a suitable host cell 1) DNA according to claims 1 to 5 or 12, and 2) a suitable reporter gene functionally linked to an operative hormone response element, said HRE being able to be activated by the DNA-binding domain of the protein encoded by said DNA;
 - b) bringing the host cell from step a) into contact with potential ligands which will possibly bind to the ligandbinding domain of the protein encoded by said DNA from step a);
 - c) monitoring the expression of the protein encoded by said reporter gene of step a).

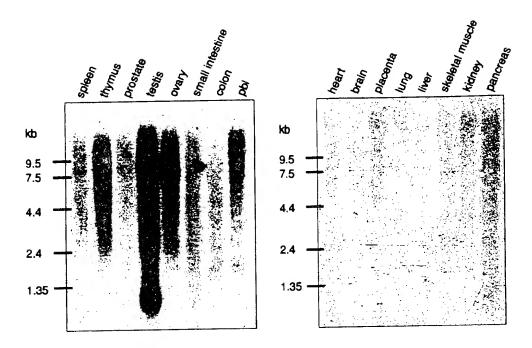
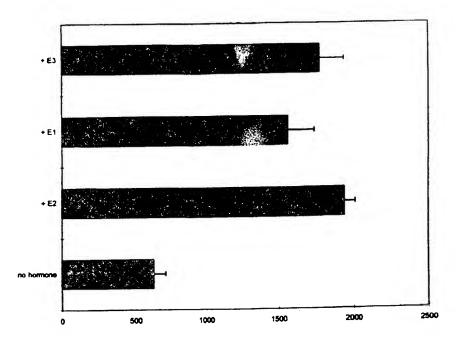


Figure 1

Allering garage States and a state



luciferase units

Fig. 2

Transient transfection of CHO cells with Estrogen Receptors Alpha and Beta Incubation with Estradiol and ICI

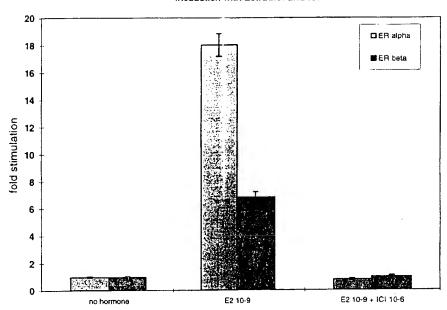


Figure 3

ERα and ERβ RT PCR on tissue-representative cell lines

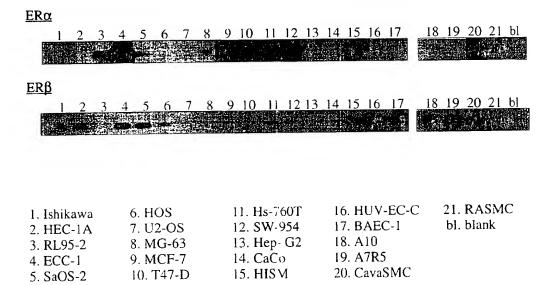
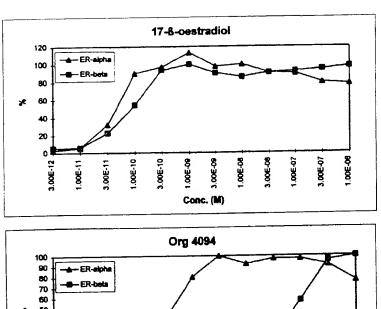
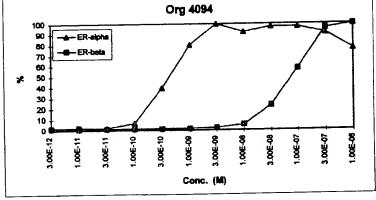


Figure 4





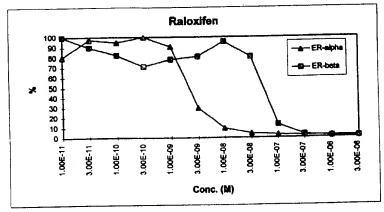


Figure 5

(12)

EUROPEAN PATENT APPLICATION

- (88) Date of publication A3: 29.12.1997 Bulletin 1997/52
- (43) Date of publication A2: 01.10.1997 Bulletin 1997/40
- (21) Application number: 97200903.9
- (22) Date of filing: 25.03.1997

- (51) Int CL⁶. **C12N 15/12**. C12N 15/62, C12N 15/85, C07K 14/72, C12N 1/21, C12N 5/16, C12Q 1/00, C12Q 1/68
- (84) Designated Contracting States:
 AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
 NL PT SE
- (30) Priority: 22.11.1996 EP 96203284 26.03.1996 EP 96200820
- (71) Applicant: Akzo Nobel N.V. 6824 BM Arnhem (NL)
- (72) Inventors:
 Mosselman, Sietse
 5346 VM OSS (NL)

- Dijkema, Rein
 5345 ML Oss (NL)
- (74) Representative: Ogilvie-Emanuelson, Claudia Maria et al Patent Department Pharma N.V. Organon P.O. Box 20 5340 BH Oss (NL)

(54) Estrogen receptor

(57) The present invention relates to isolated DNA encoding novel estrogen receptors, the proteins encod-

ed by said DNA, chimeric receptors comprising parts of said novel receptors and uses thereof.



EUROPEAN SEARCH REPORT

Application Number EP 97 20 0903

Category	Citation of document with ind of relevant passa		Relevant to claim	CLASSIFICATION OF THE APPLICATION (INLC).6)
Y,D	GREEN, G.L.: "Seque human estrogen recep DNA." SCIENCE, vol. 231, 13 March pages 1150-1154, XPC	ence and expression of otor complementary	1-14	C12N15/12 C12N15/62 C12N15/85 C07K14/72 C12N1/21 C12N5/16 C12Q1/00 C12Q1/68
ł	orphan receptors spi steroid receptor su PROCEEDINGS OF THE SCIENCE, vol. 91, June 1994, pages 6040-6044, XP! * the whole documen	NATIONAL ACADEMY OF	1-14	
A,D	EVANS, R.M.: "The hormone receptor su SCIENCE, vol. 240, 13 May 19 pages 889-895, XP00 * the whole documen	1-14	TECHNICAL FIELDS SEARCHED (Int.CI.6) C12N C07K C12Q	
A	EP 0 371 820 A (SAL STUDI) 6 June 1990 * page 4, lines 20-	K INST FOR BIOLOGICAL 27 and claim 19 *	11-14	
E	WO 97 09348 A (KARO J M (SE); ENMARK EV * the whole documen	BIO AB ;KUIPER GEORGE (A (SE); GUSTAFSSON) t *	1-14	
		-/		
	The present search report has t	peen drawn up for all claims		
	Place of search	Date of completion of the search		Examiner
	THE HAGUE	23 October 1997	Ma	ındl, B
Хра	CATEGORY OF CITED DOCUMENTS rticularly relevant if taken alone rticularly relevant if combined with another	T : theory or prince E : earlier patent di after the filing d D : document cited	ocument, but put ate	olished on, or

PO FORM 1503 03 82



eren geringen in der Gregoria

A STATE OF THE STA

EUROPEAN SEARCH REPORT

Application Number EP 97 20 0903

		RED TO BE RELEVANT	Belevant	
ategory	Citation of document with in- of relevant passa	CLASSIFICATION OF THE APPLICATION (InLCI.6)		
Y,D	human estrogen reception." SCIENCE, vol. 231, 13 March pages 1150-1154, XP	1986,	1-14	C12N15/12 C12N15/62 C12N15/85 C07K14/72 C12N1/21 C12N5/16 C12Q1/00 C12Q1/68
Y	orphan receptors sp steroid receptor su PROCEEDINGS OF THE SCIENCE, vol. 91, June 1994, pages 6040-6044, XP * the whole documen	NATIONAL ACADEMY OF	1-14	
A,D	EVANS, R.M.: "The hormone receptor su SCIENCE, vol. 240, 13 May 19 pages 889-895, XP00 * the whole documen	88, 2019515	1-14	TECHNICAL FIELDS SEARCHED (Int.CL6) C12N C07K C12Q
A	EP 0 371 820 A (SAL STUDI) 6 June 1990 * page 4, lines 20-	K INST FOR BIOLOGICAL 27 and claim 19 *	11-14	
E	WO 97 09348 A (KARO J M (SE); ENMARK EV * the whole documen	BIO AB ; KUIPER GEORGE G A (SE); GUSTAFSSON) t * 	1-14	
	The present search report has			
	Place of search	Date of completion of the search		Examiner
X pa Y∶pa	THE HAGUE CATEGORY OF CITED DOCUMENTS riscularly relevant if taken alone riscularly relevant if combined with anot pument of the same category	23 October 1997 T theory or principl E 'earlier patent do after the filips be D: document ofted L document ded I	e underlying tournent, but pite n the applicat	ublished on, or Jan
A:tec	the background the background the background the written disclosure ermediate document	& : member of the s document		



A STATE OF THE STA

EUROPEAN SEARCH REPORT

Application Number EP 97 20 0903

ategory	Citation of document with indication		Relevant	CLASSIFICATION OF THE		
at e gory	of relevant passages		to claim	APPLICATION (Int.Cl.6)		
, х	MOSSELMAN S. ET AL.: "	ER-beta:	1-10			
	Identification and char-	acterization of a				
	novel human estrogen re	ceptor."				
	FEBS LETTERS,					
	vol. 392, no. 1, 19 Aug	ust 1996,				
	pages 49-53, XP00204441	0				
	* the whole document *					
		•	1			
				1		
				TECHNICAL FIELDS SEARCHED (Int.Cl.6)		
				SEARCHED (III.CI.I.)		
			1			
	1		-			
	The present search report has been d	rawn up for all claims				
	Place of search	Date of completion of the search	1	Examiner		
	THE HAGUE	23 October 1997	Ma	ndl, B		
	ATEGORY OF CITED DOCUMENTS	T theory or princip	ole underlying the	invention		
X: particularly relevant if taken alone Y: particularly relevant if combined with another document of the same category A: technological background O non-written dissolosure P: intermediate document		E earlier patent de	ocument, but pub	, but published on, or		
		D : document cited	efter the filing date D: document cited in the application L: document cited for other reasons			
		& member of the				